

## 明細書

### データ処理装置及びデータ処理方法

#### 5 技術分野

本発明は、音声および映像の符号化データを復号し再生する技術に関し、特に、複数のデータストリームから音声および映像を連続して再生する際に、同じデータストリームの対応する音声および映像を同期して再生する技術に関する。

10

#### 背景技術

近年、デジタル技術の発達により、動画、静止画等の映像や、音声等に関するコンテンツのデータは符号化され、符号化データストリームとして光ディスクやハードディスク等の記録媒体に記録されるようになってきた。例えばMPEG規格（ISO11172またはISO13818）によれば、音声はオーディオ符号化ストリームとして、また映像はビデオ符号化ストリームとしてそれぞれ符号化される。その後、各符号化データを格納したデータパケットが時系列に並べられ、1つに多重化されて符号化データストリームが構築される。このような符号化ストリームの多重化処理は、システムエンコードと呼ばれている。システムエンコードされた多重化データストリーム（システムストリーム）は、データパケット単位で一つの伝送路上を伝送され、再生装置によって処理される。その結果、

映像および音声が再生される。

図1 (a) ~ (d) は、データストリーム10のデータ構造を示している。再生装置は、図1 (a) に示すデータストリーム10から順に、(b) および(c) に示すデータ構造を構築し、(d) に示す形態によって映像および音声を出力する。

図1 (a) は、データストリーム10のデータ構造を示す。例えば、データストリーム10はMPEG2トランSPORTストリームである。

データストリーム10は、多重化されたビデオパケット“Vn”( $n = 1, 2, \dots$ ) およびオーディオパケット“A $n$ ”( $n = 1, 2, \dots$ ) から構成されている。各パケットは、パケットヘッダと、パケットヘッダに続くペイロードから構成されている。ビデオパケットのペイロードには映像に関するデータが格納され、オーディオパケットのペイロードには映像に関するデータが格納されている。

図1 (b) は、パケット化エレメンタリストリーム(PES)11のデータ構造を示す。PES11は、データストリーム10を構成する各パケットのペイロードのデータに基づいて構築される。PES11は、複数のPESパケットから構成されており、各PESパケットはPESヘッダおよびPESペイロードから構成されている。

図1 (c) は、ビデオ／オーディオのエレメンタリストリーム(ES)の構成を示す。ビデオES12vは、ピクチャヘッダ、ビ

クチャデータ、およびピクチャの表示タイミングを規定する再生時刻情報（VPTS）等からなるデータ単位を複数含んでいる。各ピクチャデータは、そのデータのみによって、または、そのデータとその前および／または後に復号化されるピクチャデータとによって構築可能な1枚分のピクチャのデータである。オーディオES12aも同様に、ヘッダ、オーディオフレームデータおよびオーディオフレームの出力タイミングを規定する再生時刻情報（APTS）等からなるデータ単位を複数含んでいる。なお再生時刻情報（APT S、VPTS）は、MPEG2規格では33ビットで表されるデータであり、PESパケットのヘッダ（図1（b）の“PES-H”）の領域（Presentation\_Time\_Stamp）に格納されている。

図1（d）は、出力されるビデオピクチャおよびオーディオフレームを示す。ビデオピクチャ13-1および13-2等は1枚の画像（ピクチャ）であり、各々はビデオES12v内のピクチャデータに基づいて表示される。ピクチャの表示のタイミングはビデオES12v内の再生時刻情報（VPTS）によって指定され、その情報にしたがって各ピクチャの表示を切り替えることによって動画像が映像機器の画面上に表示される。オーディオフレームの出力のタイミングはオーディオES12a内の再生時刻情報（APTS）によって指定され、その情報にしたがって各オーディオフレームを出力することによって音声がスピーカから出力される。

図2は、図1（a）に示すデータストリーム10を再生すること

が可能な従来の再生装置 120 の機能プロックの構成を示す。再生装置 120 は、データストリーム 10 の各パケットを取得し、取得したパケットに基づいてビデオおよびオーディオの各エレメンタリストリームまでデコードし、その後、復元されたビデオピクチャおよびオーディオフレームを出力する。  
5

いま、再生装置 120 が 2 つのデータストリーム 1 および 2 を続けて読み出し、それぞれのビデオピクチャおよびオーディオフレームを再生する処理を考える。各データストリームは図 1 (a) のデータ構造を有する。ストリーム読み出し部 1201 が各データストリームを連続して読み出すと、再生装置 120 内部には 1 本のデータストリームが送出される。そこで以下では、この 1 本のデータストリームのデータ部分のうち、データストリーム 1 に対応するデータ部分を「第 1 区間」と称し、データストリーム 2 に対応するデータ部分を「第 2 区間」と称する。また、再生するストリームが切り替わる点を境界と呼ぶことにする。境界は第 1 区間の終点であり、  
10 第 2 区間の始点でもある。  
15

データストリームでは、音声および映像の各パケットが多重化されている。同時刻に再生されるべき音声および映像のパケットがデータストリームにおいて直列に配置されて伝送されるので、データストリームの読み出しを中断すると同時に再生すべき音声および映像の一方しか存在しない場合がある。よって、相対的に音声あるいは映像の一方の再生時間が短く、もう一方が長くなることがある。これは上述した境界の第 1 区間の終点近傍で生じる。このようなデ  
20

5

ータストリームを復号すると、第1区間の終点近傍（例えば第1区間の再生終了時刻の1秒前）では映像は全て再生されたが音声が途切れたり、または音声は全て再生されたが映像が途切れたりすることがある。また、第2区間の始点においても途中から読み出しを開始するため、映像の再生当初の音声が存在しなかったり、音声の再生当初の映像が存在しなかったりすることがある。

10

特に第1区間および第2区間の映像および音声を連続して再生するときは、本来同時に再生されるべきではない、境界前後の異なる区間の音声と映像が同時に再生されてしまうことがある。そこで再生装置120は、読み出しの対象を切り替えた時にダミーパケットを挿入している。図3(a)は、第1区間と第2区間との間に挿入されたダミーパケット1304を示す。ダミーパケット挿入部1202がデータストリーム1302の終端にダミーパケット1304を挿入し、その後データストリーム1303を結合することにより、ダミーパケット1304によって第1区間と第2区間に区分できるデータストリーム1301が得られる。

15

第1区間のデータストリーム1302、ダミーパケット1304および第2区間のデータストリーム1303は、連続してストリーム分離部1203に送られる。ストリーム分離部1203は第1区間のデータストリーム1302を受け取ると、それに含まれるオーディオパケット(A11等)と、ビデオパケット(V11、V12、V13等)とを分離し、さらに、それぞれをオーディオESおよびビデオESにまで復元(システムデコード)しながら、それぞれを

順次第1オーディオ入力バッファ1205および第1ビデオ入力バッファ1212に格納する。

そしてストリーム分離部1203がダミーパケット1304を検出すると、第1スイッチ1204を切り替えて、ストリーム分離部1203と第2のオーディオ入力バッファ1206を接続する。同時に第2のスイッチ1211を切り替え、ストリーム分離部1203と第2のビデオ入力バッファ1213を接続する。  
5

その後、ストリーム分離部1203は、第2区間のデータストリーム1303のオーディオパケット（A21等）と、ビデオパケット（V21、V22、V23等）とを分離して、それぞれをシステムデコードしながら、オーディオESおよびビデオESを第2オーディオ入力バッファ1206および第2ビデオ入力バッファ1213に格納する。  
10

オーディオ復号部1208は、第3スイッチ1207を介して第1オーディオ入力バッファ1205からオーディオESを読み出して復号化（エレメンタリデコード）する。そして、得られたオーディオのフレームデータをオーディオ出力バッファ1209に送る。  
15  
オーディオ出力部1210は、オーディオ出力バッファ1209から復号化されたオーディオフレームデータを読み出し、出力する。

一方、ビデオ復号部1215は、第4スイッチ1214を介して第1ビデオ入力バッファ1212からビデオストリームを読み出してエレメンタリデコードする。そして、得られたビデオのピクチャデータをビデオ出力バッファ1216に送る。ビデオ出力部121  
20

7は、ビデオ出力バッファ1216から復号化されたビデオピクチャデータを読み出し、出力する。

オーディオ復号部1208およびビデオ復号部1215は、AV同期制御部1218により、復号の開始や停止の制御がされる。また、オーディオ出力部1210およびビデオ出力部1217も、AV同期制御部1218によって出力の開始や停止の制御がされる。

第3スイッチ1207および第4スイッチ1214は、第1区間の各ビデオ／オーディオパケットを読み出した後はそれぞれ切り替わり、第2オーディオ入力バッファ1206とオーディオ復号部1208とを接続し、第2ビデオ入力バッファ1213とビデオ復号部1215とを接続する。その後行われる復号処理および出力処理は先の処理と同様である。

図3（b）は、第1区間のオーディオおよびビデオストリーム1305および1306、第2区間のオーディオおよびビデオストリーム1307および1308の各再生時刻の関係を示す。各ストリームは、図1（c）に記載されているエレメンタリストリーム（ES）であるとする。各ストリームを構成するオーディオフレーム／ビデオピクチャの再生時刻は、図1（c）に示すように再生時刻情報（APTSおよびVPTS）によって規定されている。

図3（b）によれば、第1区間のオーディオストリーム1305の再生終了時刻Taとビデオストリーム1306の再生終了時刻Tbとが一致していないことが理解される。また、第2区間のオーディオストリーム1307の再生開始時刻Tcとビデオストリーム1

308 の再生開始時刻  $T_d$  とが一致していないことも理解される。

例えば、日本国特開2000-36941号公報には、スキップ点の前後において動画を連続再生可能な、第1の例による再生装置が開示されている。そこで、そのような再生装置を用いて、図3  
5 (b) に示すビデオストリーム 1306 および 1308 を連続再生する処理を説明する。

図3 (b) に示すように、境界前の時刻  $T_a$  から  $T_b$  の区間では、  
10 オーディオストリーム 1305 が欠落している。そこで、オーディオ復号部 1208 は第1区間のオーディオストリームの復号完了後、一旦復号を停止する。次に、オーディオ復号部 1208 には第2オーディオ入力バッファ 1206 から第2区間のオーディオストリーム 1307 が入力される。

第2区間の時刻  $T_c$  から  $T_d$  の間は、ビデオストリーム 1308  
15 が欠落しているため、時刻  $T_c$  から時刻  $T_d$  の間のオーディオストリームを復号せずに破棄する。破棄の処理は、例えば、オーディオ復号部 1208 が第2入力バッファ 1206 上の読み出しアドレスを時刻  $T_c$  から時刻  $T_d$  に相当するデータが格納されたアドレスに移動することによって行われる。この破棄の処理は、オーディオス  
トリームを復号する処理に比べて非常に短い時間で完了するので、  
20 オーディオ復号部 1208 は、時刻  $T_d$  以降のオーディオストリームを復号開始の指示がAV同期制御部 1218 から来るのを待つ。  
一方、オーディオ復号部 1208 が時刻  $T_d$  以降の復号開始指示の待機状態に入るまでの間、ビデオ復号部 1215 は、第1区間の時

刻  $T_b$  までビデオストリームの復号および出力を行う。

時刻  $T_b$  までのビデオストリームの復号処理が完了した時点で、第 2 ビデオ入力バッファ 1 2 1 3 には境界後の時刻  $T_d$  以降のビデオストリームが格納されているとする。ビデオ復号部 1 2 1 5 は、

5 時刻  $T_b$  までのビデオストリームの復号に続けて時刻  $T_d$  以降の復号を行う。これにより、時刻  $T_b$  までの映像と時刻  $T_d$  からの映像は連続して再生される。時刻  $T_d$  のビデオストリームの復号を開始する際には、AV 同期制御部 1 2 1 8 は、待機させていたオーディオ復号部 1 2 0 8 を起動し、時刻  $T_d$  以降、オーディオストリーム 1 3 0 7 の復号を開始する。これにより、境界においてビデオストリームを連続的に再生し、かつ、音声と映像とを同期させて出力することが可能となる。

なお、日本国特開 2002-281458 号公報や日本国特開平 10-164512 号公報に記載されている技術によっても、境界前後において映像を連続して再生することができる。例えば、日本国特開 2002-281458 号公報によれば、オーディオストリーム 1 3 0 5、1 3 0 7 に付随する再生時刻情報を用いて、図 3 (b) に示す時刻  $T_c$  から  $T_d$  までの区間のオーディオストリームを破棄し、境界前後での連続再生を実現できる。なお、再生時刻情報用いると、オーディオストリームに対してビデオストリームが欠落している場合にはオーディオストリームを破棄できるため、オーディオストリームの不要な処理負荷をなくすとともに、第 2 区間のストリームをすぐに読み込むことができる。よって、境界前後における

る映像の連続再生も可能となる。

しかしながら、従来の技術では境界前後において映像を連続して再生することは可能であるが、映像と音声との同期がずれてしまうことがあった。以下、図4（a）および（b）を参照しながら具体的に説明する。

図4（a）は、2つの境界によって3つの区間が規定されたデータストリーム1401を示す。データストリーム1401には2つのダミーパケット1および2が含まれている。第1区間に対応するデータストリーム1402のオーディオパケットA11の後にダミーパケット1が挿入される。その後、第2区間に対応するデータストリーム1403が読み出される。続いて、データストリーム1403の最後のビデオパケットV22の後にダミーパケット2が挿入される。そしてその後、第3区間に対応するデータストリーム1404が読み出される。

ここで、第2区間には、ビデオパケットV21およびV22のみが存在し、オーディオパケットは存在していないことに留意されたい。これは、映像のたかだか数フレーム分に相当する短い区間が第2区間として規定され、その区間のデータストリーム1403内にはデコード可能なオーディオフレームに足るだけのオーディオパケットが存在しないことを意味している。例えば、MPEG2規格で録画されたデータストリームを、時間的に非常に短い区間を指定して編集した場合に、このような区間が生成される。

図4（b）は、第1区間のオーディオおよびビデオストリーム1

405 および 1406、第 2 区間のビデオストリーム 1407、お  
よび、第 3 区間のオーディオおよびビデオストリーム 1408 およ  
び 1409 の各再生時刻の関係を示す。図 4 (b) でも、各ストリ  
ームは、図 1 (c) に記載されているエレメンタリストリーム (E  
S) レベルにまで構築されたストリームであるとする。  
5

まず、映像の再生処理を説明すると、境界 1 の前後では、第 1 区  
間のビデオパケット V11 までのピクチャデータが第 1 ビデオ入力  
バッファ 1212 に格納され、第 2 区間のビデオパケット V21 お  
よび V22 のピクチャデータが第 2 ビデオ入力バッファ 1213 に  
10 格納される。いずれのデータもその後順次復号され、映像が連続的  
に再生される。そして、境界 2 の後は、第 3 区間のビデオストリー  
ムの格納先が、第 1 ビデオ入力バッファ 1212 に再び切り替えら  
れ、境界 1 と同様の制御で復号され、映像が連続的に出力される。

一方、音声の再生処理を説明すると、まず時刻 Taにおいて、オ  
15 ディオ復号部 1208 による復号が一旦停止され、オーディオス  
トリームの格納先が、第 1 オーディオ入力バッファ 1205 から第  
2 オーディオ入力バッファ 1206 に切り替えられる。次に、第 3  
区間のデータストリームが記録媒体 121 から読み出され、第 3 区  
間のオーディオストリームが、第 2 オーディオ入力バッファ 120  
20 6 に格納される。

従来の再生装置は、オーディオストリームを復号し音声を再生す  
る際に、再生時刻情報を利用している。第 2 区間のビデオストリー  
ム 1407 に付加された再生時刻情報および第 3 区間のビデオスト

リーム 1 4 0 9 に付加された再生時刻情報が単純に増加しているとき（特に、時刻  $T_c$  から時刻  $T_f$  までの区間の再生時刻情報の値が単調増加しているとき）は、特に問題なく処理を進めることができある。オーディオ復号部 1 2 0 8 およびオーディオ出力部 1 2 1 0 は、ビデオ復号部 1 2 1 5 およびビデオ出力部 1 2 1 7 が時刻  $T_f$  の処理をするまで待機し、時刻  $T_f$  より処理を開始し、映像に同期して音声を出力すればよい。

しかしながら、各区間のデータストリームに付加された再生時刻情報はストリーム相互間では調整されていないため、各区間の再生時刻情報の値の大小を予め定めたり、予測することはできない。よって再生時刻情報に基づいて再生を制御すると、破棄してはならないデータを破棄してしまう等の不具合が生じ、連続再生に支障を来たす。例えば、時刻  $T_f$  に出力すべきオーディオフレームの再生時刻情報の値 ( $APTS\_f$ ) が時刻  $T_c$  に出力すべきビデオピクチャの再生時刻情報の値 ( $VPTS\_c$ ) よりも小さいとすると、従来の再生装置では、第 2 区間の映像が再生される前あるいは再生中に第 3 区間のオーディオストリームが破棄されてしまう。特に、 $APTS\_f$  が  $VPTS\_c$  よりも非常に小さい場合には、第 3 区間のオーディオストリームのデータが大量に破棄され、第 3 区間の映像が再生され始ても音声が全く出力されないという不具合が発生する。

また、時刻  $T_f$  の再生時刻情報 ( $APTS\_f$ ) の値が、第 2 区間の先頭ビデオピクチャの再生時刻情報 ( $VPTS\_c$ ) の値以上

で、かつ最終ビデオピクチャの再生時刻情報（V P T S \_ d）の値以下の時には、第2区間の映像再生中に、時刻 T f から再生が開始されるべき第3区間の音声の再生が開始されるという不具合が発生する。

5 本発明の目的は、複数のデータストリームを連続的に再生する際に、音声と映像とをずれることなく同期して再生することである。

### 発明の開示

本発明によるデータ処理装置は、映像データおよび音声データを含むデータストリームから映像および音声を再生する。前記映像データおよび前記音声データの各々には再生のタイミングを示す時刻情報が付加されている。前記データ処理装置は、第1データストリームおよび第2データストリームを連続して取得するストリーム取得部と、前記第1データストリームと前記第2データストリームとが切り替わるデータ位置に、境界を示すダミーデータを挿入する挿入部と、前記ダミーデータを検出して、前記第1データストリームおよび前記第2データストリームに異なる識別情報を割り当て、各データストリームの前記映像データおよび前記音声データに前記識別情報を関連付ける解析部と、同じ識別情報が関連付けられた映像データおよび音声データの各時刻情報に基づいて、前記映像データに基づく映像および前記音声データに基づく音声の出力タイミングを制御する制御部と、前記出力タイミングで、前記映像および前記音声を出力する出力部とを備えている。

前記制御部は、前記映像データに付加された時刻情報および前記音声データに付加された時刻情報に基づいて、前記第1データストリームに関する前記映像および前記音声の各再生終了時刻を特定し、前記音声の再生終了時刻が前記映像の再生終了時刻よりも遅い時刻であるときは、前記映像の再生終了時刻から前記音声の再生終了時刻までの前記音声の出力を停止してもよい。

前記制御部は、前記映像データに付加された時刻情報および前記音声データに付加された時刻情報に基づいて、前記第2データストリームに関する前記映像および前記音声の各再生開始時刻を特定し、前記音声の再生開始時刻が前記映像の再生開始時刻よりも早い時刻であるときは、前記音声の再生開始時刻から前記映像の再生開始時刻までの前記音声の出力を停止してもよい。

前記制御部は、映像データおよび音声データに関連付けられた識別情報が異なるとき、前記映像データに基づく映像のみを先に出力させ、前記映像の再生終了後に得られた映像データに関連付けられた識別情報が前記音声データに関連付けられた識別情報と一致したとき、同じ識別情報が関連付けられた前記映像データおよび前記音声データの各時刻情報に基づいて、前記映像データに基づく映像および前記音声データに基づく音声の出力タイミングを制御してもよい。

前記ストリーム取得部は、3以上のデータストリームを連続して取得することが可能であり、前記挿入部は、連続して取得される2つのデータストリームが切り替わるデータ位置ごとに、識別情報に

対応する、単調増加または単調減少させた値を有するダミーデータを挿入してもよい。

前記制御部は、前記音声データに関連付けられた識別情報が、過去に出力された映像の映像データに関連付けられた識別情報と一致するときは、前記音声データに基づく音声の出力を停止し、現在出力の対象としている映像の映像データに関連付けられた識別情報と同じ識別情報を有する音声データに基づいて音声を出力してもよい。  
5

前記制御部は、前記音声データに関連付けられた識別情報がまだ出力されていない映像の映像データに関連付けられた識別情報と一致するときは前記音声データに基づく音声の出力を停止し、その後取得した映像データに関連付けられた識別情報と一致したときにおいて、前記音声データに基づいて音声を出力してもよい。  
10

前記データストリームは、映像データを格納したパケットおよび音声データを格納したパケットを含むパケット構造で構成されており、前記挿入部は、前記第1データストリームの最終パケットと、前記第2データストリームの先頭パケットとの間の位置に、境界を示すダミーパケットを挿入してもよい。  
15

本発明によるデータ処理方法は、映像データおよび音声データを含むデータストリームから映像および音声を再生する。前記映像データおよび前記音声データの各々には再生のタイミングを示す時刻情報が付加されている。前記データ処理方法は、第1データストリームおよび第2データストリームを連続して取得するステップと、前記第1データストリームと前記第2データストリームとが切り替  
20

わるデータ位置に、境界を示すダミーデータを挿入するステップと、前記ダミーデータを検出して、前記第1データストリームおよび前記第2データストリームに異なる識別情報を割り当て、各データストリームの前記映像データおよび前記音声データに前記識別情報を関連付けるステップと、同じ識別情報が関連付けられた映像データおよび音声データの各時刻情報に基づいて、前記映像データに基づく映像および前記音声データに基づく音声の出力タイミングを制御するステップと、前記出力タイミングで、前記映像および前記音声を出力するステップとを包含する。

10 前記制御するステップは、前記映像データに付加された時刻情報および前記音声データに付加された時刻情報に基づいて、前記第1データストリームに関する前記映像および前記音声の各再生終了時刻を特定するステップと、前記音声の再生終了時刻が前記映像の再生終了時刻よりも遅い時刻であるときは、前記映像の再生終了時刻から前記音声の再生終了時刻までの前記音声の出力を停止するステップとを包含してもよい。

前記制御するステップは、前記映像データに付加された時刻情報および前記音声データに付加された時刻情報に基づいて、前記第2データストリームに関する前記映像および前記音声の各再生開始時刻を特定するステップと、前記音声の再生開始時刻が前記映像の再生開始時刻よりも早い時刻であるときは、前記音声の再生開始時刻から前記映像の再生開始時刻までの前記音声の出力を停止するステップとを包含してもよい。

前記制御するステップは、映像データおよび音声データに関連付けられた識別情報が異なるとき、前記映像データに基づく映像のみを先に出力させるステップと、前記映像の再生終了後に得られた映像データに関連付けられた識別情報が前記音声データに関連付けられた識別情報と一致したとき、同じ識別情報が関連付けられた前記映像データおよび前記音声データの各時刻情報に基づいて、前記映像データに基づく映像および前記音声データに基づく音声の出力タイミングを制御するステップとを包含してもよい。

前記ストリームを取得するステップは、3以上のデータストリームを連続して取得し、前記挿入するステップは、連続して取得される2つのデータストリームが切り替わるデータ位置ごとに、識別情報に対応する、単調増加または単調減少させた値を有するダミーデータを挿入してもよい。

前記制御するステップは、前記音声データに関連付けられた識別情報が、過去に出力された映像の映像データに関連付けられた識別情報と一致するときは、前記音声データに基づく音声の出力を停止するステップと、現在出力の対象としている映像の映像データに関連付けられた識別情報と同じ識別情報を有する音声データに基づいて音声を出力するステップとを包含してもよい。

前記制御するステップは、前記音声データに関連付けられた識別情報がまだ出力されていない映像の映像データに関連付けられた識別情報と一致するときは前記音声データに基づく音声の出力を停止するステップと、その後取得した映像データに関連付けられた識別

情報と一致したときにおいて、前記音声データに基づいて音声を出力するステップとを包含してもよい。

前記データストリームは、映像データを格納したパケットおよび音声データを格納したパケットを含むパケット構造で構成されており、前記挿入するステップは、前記第1データストリームの最終パケットと、前記第2データストリームの先頭パケットとの間の位置に、境界を示すダミーパケットを挿入してもよい。  
5

#### 図面の簡単な説明

10 図1（a）～（d）は、データストリーム10のデータ構造を示す図である。

図2は、図1（a）に示すデータストリーム10を再生することが可能な従来の再生装置120の機能ブロックの構成を示す図である。

15 図3（a）は、第1区間と第2区間との間に挿入されたダミーパケット1304を示す図であり、図3（b）は、第1区間のオーディオおよびビデオストリーム1305および1306、第2区間のオーディオおよびビデオストリーム1307および1308の各再生時刻の関係を示す図である。

20 図4（a）は、2つの境界によって3つの区間が規定されたデータストリーム1401を示す図であり、図4（b）は、第1区間のオーディオおよびビデオストリーム1405および1406、第2区間のビデオストリーム1407、および、第3区間のオーディオ

およびビデオストリーム 1408 および 1409 の各再生時刻の関係を示す図である。

図 5 は、トランSPORTストリーム 20 のデータ構造を示す図である。

5 図 6 (a) はビデオ TS パケット 30 のデータ構造を示す図であり、図 6 (b) は、オーディオ TS パケット 31 のデータ構造を示す図である。

図 7 は、本実施形態による再生装置 100 の機能ブロックの構成を示す図である。

10 図 8 (a) ~ (c) は、それぞれ読み出された TS 1 ~ TS 3 を示す図であり、図 8 (d) は、ダミーパケットが挿入された TS 7 0 を示す図である。

15 図 9 (a) は、ダミーパケット 71 のデータ構造を示す図であり、図 9 (b) は、ダミーパケット 71 の具体的なデータ構造を示す図である。

図 10 は、第 1 の例による、境界近傍におけるオーディオストリームおよびビデオストリームの再生時刻の関係を示す図である。

図 11 は、第 2 の例による、境界近傍におけるオーディオストリームおよびビデオストリームの再生時刻の関係を示す図である。

20 図 12 は、第 3 の例による、オーディオストリームおよびビデオストリームの再生時刻の関係を示す図である。

図 13 は、複数の TS を連續して読み出す際の処理の手順を示す図である。

図14は、ストリーム解析部103の処理の手順を示す図である。

図15(a)は、ストリーム解析部103に入力されたTS70を示す図であり、図15(b)はPES80のデータ構造を示す図であり、図15(c)は、ビデオES82のデータ構造を示す図である。

図16は、フレームデータと識別情報とを関連付けしたオーディオ管理テーブルを示す図である。

図17は、音声および映像を出力するための、AV同期制御部118における前処理の手順を示す図である。

図18は、音声および映像を出力するための、AV同期制御部118における主要な処理の手順を示す図である。

図19は、音声を区間終端においてフェードアウト再生し、区間始端においてフェードイン再生するための、ゲインとオーディオストリームとの関係を示す図である。

図20は、音声を区間終端においてフェードアウト再生し、区間始端においてフェードイン再生するための、ゲインとオーディオストリームとの関係を示す図である。

### 発明を実施するための最良の形態

以下、添付の図面を参照しながら本発明のデータ処理装置を説明する。まず処理の対象であるデータストリームのデータ構造を説明し、その後、データ処理装置の実施形態である再生装置を説明する。

本実施形態では、データストリームを、MPEG2規格ISO-

13818-1で定義されているトランSPORTストリーム（以下、「TS」または「トランSPORTストリーム」と記述する）であるとして説明する。TSは、オーディオストリームおよびビデオストリームを多重化したシステムストリームの一つの形態として知られている。

図5は、トランSPORTストリーム20のデータ構造を示す。トランSPORTストリーム20は、複数のTSオブジェクトユニット(TS Object Unit; TOBU)21を含み、そのTOBU21は1以上のトランSPORTパケット(TSパケット)から構成されている。

TSパケットは、例えば、圧縮符号化されたビデオデータが格納されたビデオTSパケット(V\_TS\_P)30、(圧縮)符号化されたオーディオデータが格納されたオーディオTSパケット(A\_TS\_P)31の他、番組表(プログラム・アソシエーション・テーブル; PAT)が格納されたパケット(PAT\_TS\_P)、番組対応表(プログラム・マップ・テーブル; PMT)が格納されたパケット(PMT\_TS\_P)およびプログラム・クロック・リファレンス(PCR)が格納されたパケット(PCR\_TS\_P)等を含む。各パケットのデータ量は188バイトである。

以下、本発明の処理に関連するビデオTSパケットおよびオーディオTSパケットを説明する。他の種類のパケットのデータ構造およびそのデータに基づく機能は本発明の処理とは直接に関連しないため、それらの説明は省略する。

図6(a)はビデオTSパケット30のデータ構造を示す。ビデ

オ T S パケット 3 0 は、 4 バイトのトランSPORTパケットヘッダ  
3 0 a、 および、 1 8 4 バイトの T S ペイロード 3 0 b を有する。  
T S ペイロード 3 0 b には上述のビデオデータが格納されている。  
一方、 図 6 (b) は、 オーディオ T S パケット 3 1 のデータ構造を  
示す。 オーディオ T S パケット 3 1 も同様に、 4 バイトのトラン  
SPORTパケットヘッダ 3 1 a、 および、 1 8 4 バイトの T S ペイロ  
ード 3 1 b を有する。 T S ペイロード 3 1 b には上述のオーディオ  
データが格納されている。 T S ペイロード 3 0 b に格納されている  
ビデオデータおよび T S ペイロード 3 1 b に格納されているオーデ  
10 ィオデータは、 概ね、 図 1 (a) ~ (d) に示す関係に基づいて処  
理され、 映像および音声として再生される。

上述の例から理解されるように、 一般に T S パケットは 4 バイト  
のトランSPORTパケットヘッダと、 1 8 4 バイトのデータ領域と  
から構成されている。 パケットヘッダには、 そのパケットの種類を  
特定するパケット識別子 (Packet ID ; P I D) が記述されている。  
15 例えば、 ビデオ T S パケットの P I D は “0 x 0 0 2 0” であり、  
オーディオ T S パケットの P I D は “0 x 0 0 2 1” である。 デー  
タ領域は、 ビデオデータ、 オーディオデータ等のコンテンツのデー  
タや、 再生を制御するための制御データ等を格納する。 どのような  
20 データが格納されているかは、 パケットの種類に応じて異なる。 な  
お、 パケットヘッダと T S ペイロードとの間に「アダプテーション  
フィールド」と呼ばれる領域が挿入され、 制御データの伝送、 ペイ  
ロードのデータサイズの調整等に用いられる場合がある。 しかし、

本実施形態による処理の主要な特徴は、TSパケットのペイロードを利用した処理にあるため、アダプテーションフィールドが存在しない場合を例として説明する。

なお、図5、図6(a)および図6(b)はトランスポートストリームに関するデータ構造の例であるが、このデータ構造は、プログラムストリームにおける「パック」にも同様に適用できる。パックでは、パケットヘッダに続けてデータが配置されているからである。パケットヘッダの前にはパックヘッダが付加され、パックのデータ量は2048キロバイトである等の点においてパケットと相違している。なお、「パック」はパケットの1つの例示的な形態として知られている。なお、プログラムストリームの他、上述のパケット構造と同様のパケット構造を有するデータストリームであれば、以下に説明する処理は適用可能である。

図7は、本実施形態による再生装置100の機能ブロックの構成を示す。再生装置100には、Blu-ray Disc等の光ディスク120が装填されており、光ディスク120に記録されたトランスポートストリーム(TS)を読み出す。そして、TSを構成するビデオパケットおよびオーディオパケットからビデオデータおよびオーディオデータを抽出し、映像および音声を再生する。

本実施形態では、光ディスク120には複数のTS(例えば、TS1およびTS2)が記録されているとする。TSには複数のコンテンツのデータが重畠され得るが、説明の便宜上、1つのTSには1つのコンテンツが含まれているとする。なお、「コンテンツを再

生する」とは、そのコンテンツに含まれる映像および音声の各々を同期させながら再生することをいうとする。

再生処理の制御機能に関して、再生装置100は、ストリーム読み出し部101と、ダミーパケット挿入部102と、ストリーム解析部103と、AV同期制御部118と、識別情報格納部119とを有する。

またオーディオの再生処理機能に関して、再生装置100は、第1スイッチ104と、第1オーディオ入力バッファ105と、第2オーディオ入力バッファ106と、第3スイッチ107と、オーディオ復号部108と、オーディオ出力バッファ109と、オーディオ出力部110とを有する。

さらに、ビデオの再生処理機能に関して、再生装置100は、第2スイッチ111と、第1ビデオ入力バッファ112と、第2ビデオ入力バッファ113と、第4スイッチ114と、ビデオ復号部115と、ビデオ出力バッファ116と、ビデオ出力部117とを有する。

以下、個々の構成要素の機能（動作）を説明する前に、再生装置100による再生処理の概要を説明する。ストリーム読み出し部101は、光ディスク120から複数のTSを連続的に読み出し、ダミーパケット挿入部102に送る。図8(a)は読み出されたTS1を示し、図8(b)はTS1の後に読み出されたTS2を示し、図8(c)はTS2の後に読み出されたTS3を示す。このような連続的な読み出しは、例えばユーザが、プレイリスト等によって複

数の TS の各一部を再生区間として指定し、再生するときを想定している。または、ユーザが、複数の TS の各一部をつなぎ合わせる編集作業を行うためにその各一部を再生区間として指定し、再生するときを想定している。

5 ダミーパケット挿入部 102 は、 TS 1 の最終パケットの後で、かつ TS 2 の先頭パケットの前にダミーパケットを挿入する。その後、またダミーパケット挿入部 102 は、 TS 2 の最終パケットの後で、かつ TS 3 の先頭パケットの前にもダミーパケットを挿入する。図 8 (d) は、ダミーパケットが挿入された TS 70 を示す。  
10 TS 1 と TS 2 とが切り替わる位置にはダミーパケット 71-1 が挿入され、 TS 2 と TS 3 とが切り替わる位置にはダミーパケット 71-2 が挿入される。

本明細書では、ストリームが切り替わる位置をストリームの境界と呼ぶ。例えば図 8 (d) では、ダミーパケット 71-1 および 7  
15 1-2 を境界として TS 1 および TS 2 、 TS 2 および TS 3 が結合され、一つのトランスポートストリーム 70 が構成されている。この TS 70 のうち、 TS 1 に相当する部分を「第 1 区間」、 TS 2 に相当する部分を「第 2 区間」、 TS 3 に相当する部分を「第 3 区間」と呼ぶ。

20 図 9 (a) は、ダミーパケット 71 のデータ構造を示す。ダミーパケット 71 は、 4 バイトのトランスポートパケットヘッダ 71a 、および、 184 バイトの TS ペイロード 71b を有する。図 9 (b) は、ダミーパケット 71 の具体的なデータ構造を示す。トラ

ンスポートパケットヘッダ 71a には、図 5 に示す各種の TS パケットとは異なるパケット ID（例えば 0x1FFF）91 が割り当てられている。また、このトランSPORTパケットヘッダ 71a には、本実施形態による再生処理に利用される識別情報 92 が記述されている。<sup>5</sup> 識別情報 92 を利用した具体的な処理は後述する。一方、TS ペイロード 71b には、本実施形態によるダミーパケットであることを示す固有の情報（図 9 (b) では例えば “D u m m y 判別情報”）が記述されている。ダミーパケットのパケット ID およびこの固有の情報によって、他のパケットからは容易に識別される。

<sup>10</sup> 再び図 7 を参照する。ストリーム解析部 103 は、結合された TS 70 のうち、はじめに受け取る TS 1 のオーディオ TS パケットとビデオ TS パケットとをそれぞれ分離する。さらにストリーム解析部 103 は、分離したオーディオ TS パケットをシステムデコードしてオーディオエレメンタリストリーム（以下「オーディオストリーム」）を生成し、分離したビデオ TS パケットをシステムデコードしてビデオエレメンタリストリーム（以下「ビデオストリーム」）を生成する。<sup>15</sup> このときストリーム解析部 103 は、最初に読み出された TS 1 にデフォルトの識別情報を割り当てる。識別情報として、TS 1 を特定するための一意の値が割り当たられる。  
<sup>20</sup> そして、その識別情報を、システムデコードしたオーディオストリームとビデオストリームとに関連付ける。

ここで「関連付ける」とは、オーディオストリームに関しては、オーディオストリームのフレーム先頭アドレス値、そのフレームの

再生時刻情報（A P T S）および識別情報を対応させたテーブルを生成することをいう（例えば後述の図16）。ビデオストリームに関しては、ビデオストリーム内のピクチャヘッダの後（ピクチャデータの前）に、識別情報を記述することをいう（例えば後述の図1  
5 5（c））。

ストリーム解析部103は、上述の処理をしたオーディオストリームおよびビデオストリームを、それぞれ第1スイッチ104および第2スイッチ111を介して第1オーディオ入力バッファ105および第1ビデオ入力バッファ112に送る。

TSが順次読み出されると、ストリーム解析部103は、結合されたTS70内のダミーパケットを検出し、ダミーパケットに記述された識別情報を取得する。ダミーパケット中の識別情報の値はそのデフォルトの識別情報の値とは異なっており、その識別情報が示す値をTS2の識別情報として割り当てる。そして、TS1のオーディオパケットおよびビデオパケットに対して行った処理と同様の処理を、TS2のオーディオパケットおよびビデオパケットに対して行う。TS2のオーディオパケットおよびビデオパケットから構築されたオーディオストリームおよびビデオストリームは、それぞれ第2オーディオ入力バッファ106および第3スイッチ107に送られる。  
10  
15  
20

オーディオおよびビデオの各バッファ105、106、112および113へのストリームの蓄積と並行して、オーディオ復号部108およびビデオ復号部115は、オーディオストリームおよびビ

デオストリームを復号（エレメンタリデコード）し、再生出力可能なピクチャデータおよびオーディオフレームを生成する。

A V 同期制御部 118 は、復号後に得られたピクチャデータおよびオーディオフレームと、復号前のオーディオストリームおよびビデオストリームに関連付けられていた識別情報との対応関係を管理している。そして A V 同期制御部 118 は、その対応関係に基づいて、音声および映像を出力するタイミングを制御する。具体的には A V 同期制御部 118 は、同じ識別情報と関連付けされているピクチャデータとオーディオフレームとを再生出力の対象とする。同じ TS から得られた音声および映像の各データには同じ識別情報が付加されているため、各バッファ 105、106、112 および 113 から異なる区間を構成する TS 1～3 のどのパケットが読み出されても、再生の対象とする TS の音声および映像のみを確実に特定できる。さらに、再生時刻情報に基づいてその音声および映像の出力のタイミングを制御するため、TS 生成時に想定されていたとおりの音声および映像の同期再生が実現される。

ここで、図 10 から図 13 を参照しながら、再生装置 100 において行われる音声および映像の同期再生の種々の態様を説明する。本実施形態では、再生装置 100 は境界の前後の映像が連続的に表示されるように再生を制御する。

図 10 は、第 1 の例による、境界近傍におけるオーディオストリームおよびビデオストリームの再生時刻の関係を示す。オーディオストリームおよびビデオストリームはそれぞれエレメンタリストリ

ーム（例えば図1（c））であるとし、各ピクチャデータ等に対応して設けられた再生時刻情報PTSによって指定された時刻が再生時刻として記載されている。これは、後述の図11および12でも同じである。

5 第1区間では、ビデオストリーム502の映像の再生終了時刻Ta  
aは、オーディオストリーム501の音声の再生終了時刻Tbよりも早い。再生終了時刻Taに至るまでは、再生装置100は再生時刻情報PTSに基づいて、音声および映像を同期して再生する。一方、時刻Ta以降、音声の再生を継続すると、ビデオストリームが  
10 もはや存在しないので映像が途切れることになる。よって、第2区間の映像に連続的に接続するためには、音声の再生は不要であると判断する。そこで再生装置100は、時刻Ta以降、時刻Tbまでのオーディオストリーム501を破棄する。

なお本実施形態では、第1区間の映像の再生が終了した時刻Ta  
15 以降は、第1区間の音声と第2区間の映像とが同時に再生され出力されることはない。その理由は、再生装置100は異なる識別情報が付加された音声および映像の同時再生を禁止しているからである。これは、上述のように各区間を規定するTSに異なる識別情報を割り当てているためである。これにより、異なる区間のTSに由来する映像と音声とが同時に再生されることがなくなる。

一方、第2区間の再生を説明すると、ビデオストリーム504の映像の再生開始時刻Tcはオーディオストリーム503の音声の再生開始時刻Tdよりも早い。このとき再生装置100は、時刻Tc

から時刻  $T_d$  まで間は、映像を再生して出力し、音声の再生を停止する。時刻  $T_d$  以降は音声および映像を同期して再生する。なお、第 1 区間の映像と第 2 区間の映像は途切れることなく再生される。

なお、境界を基準としたとき、オーディオストリーム 501 およびビデオストリーム 502 の再生はそれぞれ時刻  $\Delta T_x$  および  $\Delta T_y$  だけ早く終了する。これは、1 オーディオフレームや 1 ピクチャのデータが複数のパケットに分割して格納されていることに起因する。例えば、1 ピクチャのデータを格納した全てのパケットが揃う前に  $T_S$  が切り替えられた場合には、その直前まで完成したピクチャデータしか処理できない。よって、境界の直前まで再生可能なデータが存在しない場合がある。

図 11 は、第 2 の例による、境界近傍におけるオーディオストリームおよびビデオストリームの再生時刻の関係を示す。第 1 区間では、オーディオストリーム 601 の音声の再生終了時刻  $T_a$  は、ビデオストリーム 602 の映像の再生終了時刻  $T_b$  よりも早い。このとき、再生装置 100 は音声の再生が終了した時刻  $T_a$  以降も再生終了時刻  $T_b$  まで映像を再生する。一方、第 2 区間では、音声の再生開始時刻  $T_c$  は映像の再生開始時刻  $T_d$  よりも早い。このとき再生装置 100 は、オーディオストリーム 603 の時刻  $T_c$  以降、時刻  $T_d$  までのデータを破棄する。時刻  $T_d$  以降は再生装置 100 の音声および映像を同期して再生する。なお図 11 に示す例でも、識別情報をを利用して第 1 区間の映像と第 2 区間の音声とが同時に再生されることはない。一方、第 1 区間の映像と第 2 区間の映像は途切

れることなく再生される。

図12は、第3の例による、オーディオストリームおよびビデオストリームの再生時刻の関係を示す。この例は、図8(a)～(c)に示すTS1～TS3が連続して読み出されたときに対応する。第2区間には、音声として再生可能なオーディオデータが存在しないとする。図12には、第2区間にはオーディオストリームが存在していないことが理解される。

図12の第1区間の各ストリームに関しては、図11の第1区間の処理と同じ処理によって時刻Tbまで映像が再生される。次に、第2区間のビデオストリーム703の映像が再生される。このとき再生装置100のオーディオ入力バッファにはすでに第3区間のオーディオストリーム704が存在している場合がある。しかし、第2区間のTS2および第3区間のTS3には異なる識別情報が割り当てられているため、この識別情報を利用すると第2区間の映像の再生中に第3区間のオーディオストリーム704の音声が再生されることはない。具体的には、再生装置100は第2区間の映像再生中には異なる識別情報が割り当てられたオーディオストリーム704の再生を待機させる。その後、第2区間の再生が終了すると、図10の第2区間と同じ処理により、映像の再生のみが時刻TeからTfまで行われる。時刻Tf以後、再生装置100は、第3区間の音声の待機を解除して音声および映像を同期再生する。

上述のように、再生装置100は、各区間のTSから得られた映像および音声の再生タイミングを、識別情報をを利用して制御してい

る。識別情報を利用することにより、ある TS の音声の再生時刻情報 (APTS) と他の TS の映像の再生時刻情報 (VPTS) とが一致していても同時に再生されることはない。すなわち、ストリームの再生時刻情報 (PTS) の影響を受けることなく、同じ TS から得られたオーディオストリームおよびビデオストリームのみを同期再生することができる。  
5

次に、上述の再生処理を実現するための、再生装置 100 の各構成要素の具体的な処理を説明する。

図 13 は、複数の TS を連続して読み出す際の処理の手順を示す。  
10 この処理は、ストリーム読み出し部 101 およびダミーパケット挿入部 102 によって行われる。

まず、ストリーム読み出し部 101 は、ステップ S201において光ディスク 120 からの TS の読み出しを開始するよう指示されると、ステップ S202において、ストリームの境界を判別するための識別情報 n を初期化してデフォルト値（例えば 0）を設定する。  
15 本実施形態では、識別情報 n は初期値 0 から単調に増加する整数であるとして説明するが、識別情報のデフォルト値（初期値）を一定値（例えば 100）に設定し、単調に減少する整数としてもよい。  
次に、ステップ S203において、ストリーム読み出し部 101 は  
20 第 1 区間に応するシステムストリーム (TS1) を TS パケット単位で読み出し、ダミーパケット挿入部 102 へ送り、ステップ S204 に進む。

ステップ S204において、ダミーパケット挿入部 102 は、T

Sが切り替えられて、新たな区間であるか否かを判定する。このとき、TS1のTSパケットの受信時はステップS207に進み、その後、読み出しの対象がTS1からTS2に切り替えられたときはステップS205に進む。読み出し対象のTSが切り替えられたという情報は、例えばCPU(図示せず)によって発せられる。CPUは、再生開始時点で予め特定されている再生経路に基づいてどのタイミングでTSを切り替えればよいか認識している。  
5

ステップS205では識別情報nの値に1が加算される。ステップS206では、ダミーパケット挿入部102は識別情報nを含むダミーパケットを生成し、TS1の最終パケットに付加する。ステップS207では、ダミーパケット挿入部102はTSパケットをストリーム解析部103に送出する。そして、ストリーム読み出し部101およびダミーパケット挿入部102は、対象となる全てのTSのTSパケットの送出が終了するまで、ステップS203からの処理を繰り返す。  
10  
15

ここで、再び図9(b)を参照しながら、ダミーパケット71のデータ構造を詳細に説明する。図9(b)に示すダミーパケット71は、MPEG規格として規定されているTSのNullパケットの構造を利用して構成されている。このダミーパケット71は、本実施形態において定義される符号列であって、各TSには存在し得ない符号列である。  
20

ダミーパケット挿入部102は、以下のようにしてダミーパケット71を生成する。すなわちダミーパケット挿入部102は、まず

M P E G 規格における N u l l パケットと同様に、ダミーパケット 7 1 の P I D を “ 0 x 1 F F F ” に設定する。さらにダミーパケット挿入部 1 0 2 は、ダミーパケット 7 1 内に本実施形態による固有の情報を設定する。具体的には、ダミーパケット挿入部 1 0 2 は、  
5 “ c o n t i n u i t y \_ c o u n t e r ” フィールド 9 2 に識別情報 n ( 0 x 0 ~ 0 x F ) の値を記述する。 T S 1 の先頭にはダミーパケット 7 1 が付加されていないため、 T S 1 と T S 2 との境界に挿入される最初のダミーパケット 7 1 には、識別情報 n として”  
1 1 “ が設定される。なお、識別情報は、 ” S t u f f i n g \_ d a  
10 t a ” フィールドや ” R e s e r v e d ” フィールドに記述されてもよい。

また、ダミーパケット挿入部 1 0 2 は、 “ P a y l o a d \_ u n  
i t \_ s t a r t \_ i n d i c a t o r ” フィールドに、例えば  
15 “ 1 ” を設定する。このフィールドは、 M P E G 規格に準拠した N  
u l l パケットでは “ 0 ” とされている。またダミーパケット挿入部 1 0 2 は、新たに “ D u m m y 判別情報 ” フィールドを設け、例  
えれば文字列 “ D U M ” を格納する。また、新たに “ D u m m y \_ I  
D ” フィールドを設けて、例えば、“ 0 x F ” として T S の境界を表すことにしておく。これは、他の目的で別の T S パケットを将来定義するための手当てである。以上の設定により、ダミーパケット 7 1 を検出して解析することにより、後述のストリーム解析部 1 0  
20 3 は、このダミーパケット 7 1 において、 T S の境界であると判定できる。

なお本実施形態では、ダミーパケット 71 には、後続の TS のオーディオ再生先頭時刻情報 (audio\_start PTS) 、ビデオ再生先頭時刻情報 (video\_start PTS) 、オーディオ再生終端時刻情報 (audio\_end PTS) 、および、ビデオ再生終端時刻情報 (video\_end PTS) が格納される。これらは、再生の対象となる各 TS が特定された時点で取得することができるため、予めそれらの情報を読み出しておけばよい。各時刻情報は各区間の先頭または終端近傍において、音声のフェードインおよびフェードアウトの制御をする際に利用できる。

フェードインおよびフェードアウトの制御処理は後述する。

図 14 は、ストリーム解析部 103 の処理の手順を示す。ストリーム解析部 103 は、ステップ S209 においてダミーパケット挿入部 102 から TS パケットを受け取って各パケットを解析し、ステップ S210 に進む。S210 では、ストリーム解析部 103 は、入力された TS パケットがダミーパケット 71 であるか否かを判定する。ダミーパケットであればステップ S211 に進み、そうでなければステップ S213 に進む。

ステップ S211 では、ストリーム解析部 103 はダミーパケット 71 から識別情報の値 n を抽出し、ステップ S212 において、次のダミーパケット 71 が入力されるまで一時的に記憶しておく。なお、最初のダミーパケット 71 が検出されるまでは、デフォルト値 “0” が識別情報の値として記憶されている。ダミーパケットを検出すると、第 1 スイッチ 104 および第 2 スイッチ 111 が一方

から他方に切り替えられる。

ステップ S 213 では、ストリーム解析部 103 は入力された TS パケットがビデオパケットか否かを判定する。ビデオパケットであればステップ S 214 に進み、そうでなければステップ S 216 に進む。ステップ S 214 では、ストリーム解析部 103 はビデオパケットをシステムデコードし、ビデオストリームを出力する。このとき、ビデオストリーム内のピクチャデータに識別情報を関連付ける。その後ストリーム解析部 103 は、ステップ S 215において、ビデオストリームを第 1 ビデオ入力バッファ 112 または第 2 ビデオ入力バッファ 113 に出力される。

次に、図 15 (a) ~ (c) を参照しながら、識別情報の関連付けに関する処理を具体的に説明する。図 15 (a) は、ストリーム解析部 103 に入力された TS 70 を示す。ここでは、ビデオパケット 70a ~ 70e およびダミーパケット 71 を示している。

ストリーム解析部 103 は、例えば図 15 (a) に示す TS パケット 70a から TS パケットヘッダ 70a-1 を除去して TS ペイロード 70a-2 を取得し、パケット化エレメンタリストリーム (PES) 80 を生成する。図 15 (b) は PES 80 のデータ構造を示す。PES 80 は、複数の PES パケット 81a、81b 等から構成されている。PES パケット 81a は PES ヘッダ 81a-1 および PES ペイロード 81a-2 から構成されている。

ストリーム解析部 103 は PES 80 の PES ヘッダ 81a を解析して、PES ペイロード 81a-2 内のピクチャデータの再生時

刻情報（PTS）を含むか否かを判定する。例えば、PESヘッダ81aには、PTSが記述されているか否かを示すフラグのフィールドが設けられており、そのフラグの値に基づいてPTSを含むか否かを判定できる。PTSを含む場合には、そのPTSはPESヘッダ内に格納されている。本明細書ではPTSを含むとして説明する。

PESペイロード81a-2等には各ピクチャのエレメンタリレベルのビデオストリーム（ES）82を構成するデータが格納されている。ストリーム解析部103は、PESペイロード81a-2等に基づいて、ES82を生成する。図15(c)は、ビデオES82のデータ構造を示す。ストリーム解析部103はビデオES82の生成するに際し、ピクチャヘッダ82aおよびピクチャデータ82bとの間に、VPTS82cおよびそのTSに割り当てられた識別情報82dを記述する。これにより、ピクチャデータと識別情報とが関連付けられる。なお、図15(a)に示すビデオTSパケット70eは、ダミーパケット71後に存在するため、ビデオTSパケット70a～70dとは異なるTSを構成している。よって、図15(c)に示すように、ビデオTSパケット70eから得られる識別情報82jの値(n+1)は、ダミーパケット前の各TSパケットから得られる識別情報82d、82hの値(n)よりも1つインクリメントされて関連付けられている。

図15(a)および(b)では、TSパケットと、PESパケット81bのPESペイロード部分のデータとの対応を示していない

が、これは記載の便宜のためである。実際には、そのデータは、ビデオTSパケット70dとダミーパケット71との間に存在するビデオTSパケット（図示せず）のTSペイロードに基づいて構築される。また、ビデオTSパケット70eは、そのTSにおいてダミーパケット71以降で最初のPESヘッダを含むTSパケットであるとしている。したがって、ダミーパケット71とビデオTSパケット70eとの間にはPESヘッダを含まないTSパケットが存在し得る。

次に、図14のステップS216以降の処理を説明する。ステップS216では、ストリーム解析部103は入力されたTSパケットがオーディオパケットか否かを判定する。オーディオパケットであればステップS217に進み、そうでなければステップS210に戻る。

ステップS217では、ストリーム解析部103はオーディオパケットをシステムデコードし、オーディオストリームを出力する。このとき、オーディオストリーム内のフレームデータと識別情報とを関連付ける。その後ストリーム解析部103は、ステップS218において、オーディオストリームを第1オーディオ入力バッファ105または第2オーディオ入力バッファ106に出力する。

図16は、フレームデータと識別情報とを関連付けしたオーディオ管理テーブルを示す。オーディオ管理テーブルには、識別情報と、再生対象のオーディオフレームの再生時刻(APTS)と、オーディオフレームのデータを格納したオーディオ入力バッファ105ま

5

たは 106 のアドレスとが対応付けられて記述されている。ストリーム解析部 103 は、生成したテーブルを識別情報格納部 119 に送る。識別情報格納部 119 はそのテーブルを保持する。なお、図 16 の例では、第 1 区間が 2 フレーム、第 2 区間が 3 フレームのみで構成されているが、これは識別情報 n が変化するということを説明するための例であり、一般にフレーム数はより多い場合が多い。

10

オーディオ復号部 108 は、オーディオ入力バッファ 105 および 106 に格納されたオーディオストリームを読み出す。例えば、オーディオ復号部 108 は、識別情報格納部 119 に格納されたテーブル（図 16）を参照して、そのテーブルのエントリ単位でオーディオストリームを読み出す。その後、オーディオ復号部 108 はそのオーディオストリームに対してエレメンタリデコードを行う。その結果、圧縮符号化が解除されたオーディオフレームデータが得られる。オーディオ復号部 108 はそのオーディオフレームデータをオーディオ出力バッファ 109 に格納する。格納先のアドレスは AV 同期制御部 118 によって指定される。

15

ビデオ復号部 115 は、ビデオ入力バッファ 112 および 113 に格納されたビデオストリームを読み出して、エレメンタリデコードを行う。その結果得られたピクチャデータもまた、圧縮符号化が解除されている。ビデオ復号部 115 はそのビデオピクチャデータをビデオ出力バッファ 116 に格納する。格納先のアドレスは AV 同期制御部 118 によって指定される。

AV 同期制御部 118 は、ビデオ出力バッファ 116 のどの位置

(アドレス) にビデオピクチャデータを格納するかをビデオ復号部 115 に指示する。そして、格納したビデオピクチャデータを再生するための情報を収集してビデオ管理テーブルを生成する。ビデオ管理テーブルは、識別情報と VPTS と格納アドレスとを対応付けて構成されており、図 16 に示すオーディオ管理テーブルにおいて “APTS” を “VPTS” に読み替えればよい。ただし、ビデオ出力バッファ 116 のバッファ容量は、例えば 3 ~ 4 フレームデータを格納する程度であるため、図 16 に示すような、識別情報と VPTS と格納アドレスとの組からなるエントリを 7 以上も設ける必要はない。AV 同期制御部 118 は、ビデオ管理テーブルを識別情報格納部 119 に格納してもよいし、自身の内部バッファ（図示せず）に保持してもよい。

次に、図 17 および図 18 を参照しながら、AV 同期制御部 118 の音声および映像の出力制御処理を説明する。以下の説明では、オーディオストリームに関連付けられた識別情報（図 16）を “na” と記述し、ビデオストリームに関連付けられた識別情報（図 15(c)）を “nv” と記述する。識別情報 na はオーディオ管理テーブル（図 16）から取得される。識別情報 nv は AV 同期制御部 118 によって生成されたビデオ管理テーブルから取得される。なお、識別情報 na および nv は、対応する各再生時刻情報 APTS および VPTS よりも充分早く検出され、指定された再生時刻までに AV 同期の処理や、復号した信号を破棄するか出力するかの判定を実行できるとする。また、識別情報 na および nv は、ストリ

ーム解析部 103 によって、各区間に含まれるオーディオフレームおよびビデオフレームのそれぞれの先頭フレームに必ず付与され、識別情報格納部 119 に格納されているとする。

図 17 は、音声および映像を出力するための、AV 同期制御部 118 における前処理の手順を示す。AV 同期制御部 118 は、ステップ S306 において、次に出力しようとするオーディオストリームの識別情報  $n_a$  を取得し、ステップ S307 において、次に出力しようとするビデオストリームの識別情報  $n_v$  を取得する。その後の処理は図 18 に示すボックス「A」に進む。一方、ステップ S306 および S307 へ至るボックス「B」および「C」は図 18 に示すボックス「B」および「C」から継続される手順である。

図 18 は、音声および映像を出力するための、AV 同期制御部 118 における主要な処理の手順を示す。ステップ S401 において、AV 同期制御部 118 は、これから出力しようとしているオーディオフレームが、第 1 区間と第 2 区間の境界直後であるかどうか、すなわち、第 2 区間の先頭のオーディオフレームであるかどうかを判定する。この判定は、オーディオストリームの識別情報  $n_a$  と、1 制御単位前のオーディオストリームの識別情報が一致するか否かを比較することによって行われる。ここで、1 制御単位とは、オーディオ復号部 108 がオーディオストリームを復号する際のオーディオフレーム単位、または AV 同期制御部 118 が、AV 同期を行う時間間隔の単位のことであり、例えば、ビデオのフレーム単位あるいはフィールド単位と同期しているとする。

これから出力しようとしているオーディオフレームが境界後の先頭フレームではない場合には、処理はステップ S 4 0 2 に進み、境界後の先頭フレームである場合には処理はステップ S 4 0 6 に進む。ステップ S 4 0 2 では、AV同期制御部 1 1 8 は、これから出力しようとしているオーディオフレームの再生終了時刻が、その第 1 区間のビデオピクチャの再生終了時刻よりも遅い時刻まで再生するように指定されているか否かを判定する。指定されていない場合にはステップ S 4 0 3 に進み、指定されている場合にはステップ S 4 0 5 に進む。図 1 1 に示す例では、オーディオフレームの再生時刻はビデオピクチャの第 1 区間の終了時刻よりも早いため、このようなときには処理はステップ S 4 0 3 に進む。一方、図 1 0 に示す例では、オーディオフレームの再生時刻はビデオピクチャの第 1 区間の終了時刻よりも遅いため、処理はステップ S 4 0 5 に進む。

なお、ステップ S 4 0 2 の判定は、オーディオストリーム内に記述された再生時刻情報 A P T S の値と、ビデオストリーム内に記述された再生時刻情報 V P T S の値とを比較することによって行われる。A P T S の値が V P T S の値よりも大きい場合には、その A P T S の値に対応するオーディオフレームはその V P T S の値に対応するビデオピクチャよりも遅い時刻に（すなわち後に）再生され、逆に小さい場合には、そのオーディオフレームはそのビデオピクチャよりも早い時刻に（すなわち先に）再生される。この判定手法は、以下の再生時間の比較でも同様に利用される。

なおトランSPORTストリームでは、システムクロックのいわゆ

るラップアラウンドが許可されている。ラップアラウンドとは、システムクロックがある一定値に達するとその後は0からカウントが開始されることをいう。本明細書では、上述の再生時刻情報の値の比較に際してはラップアラウンドがないとして説明する。ラップアラウンドが行われた場合には、システムクロック値は一旦 $0 \times 00000000$ になる。その値をはさむときは、再生時刻情報の値が大きい方が早い時刻に再生され、小さいほうが後の時刻に再生されることを意味する。

ステップS403では、AV同期制御部118はAV同期処理を行う。AV同期処理とは、第1区間の音声の再生時刻情報APTSと映像の再生時刻情報VPTSとをそれぞれ再生装置の基準時計と比較し、基準時計に対して再生すべき音声および／または映像が進んでいれば、AV同期制御部118がオーディオ出力部110および／またはビデオ出力部117に指示を出して出力を遅延させる。  
逆に、遅れていればスキップ処理を指示し、出力される映像および音声の出力時刻を調整する。なお、基準時計は、TSに含まれる基準時計の情報、または、APTSやVPTS自身のいずれか一つを基準として計時を行ってもよい。ステップS404では、オーディオ出力部110は、AV同期制御部118の指示を受けてオーディオ出力バッファ109に格納されているオーディオフレームデータを出力し、また、ビデオ出力部117はAV同期処理部118の指示を受けて、ビデオ出力バッファ116に格納されているビデオピクチャデータを出力する。これによりユーザは再生された映像およ

び音声を視聴できる。

ステップ S 405 では、AV 同期制御部 118 はオーディオフレームのデータの一部を破棄する。対象は、ビデオピクチャの第 1 区間の終了時刻以降の第 1 区間のオーディオストリームのデータである。ここで、オーディオフレームを破棄することは、出力バッファ 109 内に格納されている第 1 区間のオーディオフレームを消去し、または無視することによって実現される。オーディオフレームを破棄した後は、処理はステップ S 306 (図 17) に戻る。この期間中は映像はビデオ出力部 117 から出力される。そして、AV 同期制御部 118 は、第 2 区間の映像および音声に関する識別情報（境界後の先頭の識別情報）に基づいて処理を改めて実行する。第 2 区間の再生制御は後述する。

なお、上述のステップ S 402 の判定およびその判定結果に基づく処理は、オーディオ出力バッファ 109 中のオーディオデータに対して行うとして説明したが、入力バッファ 105 および 106 に格納されているオーディオストリームに対して行っててもよい。特に、ストリームを破棄する場合には入力バッファに格納されているオーディオストリームを読み出すポインタを識別情報格納部 119 に格納している第 2 区間の先頭アドレスまで移動するだけでよいので、処理が簡単になる。

ステップ S 406 では、オーディオ出力部 110 はオーディオフレームデータの出力を一旦停止する。ステップ S 401 からステップ S 406 への分岐は、出力しようとしているオーディオフレーム

データが T S 7 0 における新たな区間（ここでは第 2 区間）の先頭のフレームデータであることを意味する。

ステップ S 4 0 7 では、A V 同期制御部 1 1 8 は、オーディオストリームの識別情報 n a とビデオストリームの識別情報 n v とが一致するか否かを比較する。一致する場合は、現在処理しているビデオストリームおよびオーディオストリームが同じ T S に格納されていたことを意味している。一致する場合には処理はステップ S 4 0 8 に進み、一致しない場合には処理はステップ S 4 1 0 に進む。

ステップ S 4 0 8 では、A V 同期制御部 1 1 8 は、これから出力しようとしているオーディオフレームの再生開始時刻がビデオピクチャの再生開始時刻よりも遅いか否かを判定する。遅い場合（V P T S の値 < A P T S の値の場合）にはステップ S 4 0 3 に進み、早い場合（V P T S の値 > A P T S の値の場合）にはステップ S 4 0 9 に進む。図 1 1 に示す例では、第 2 区間においてオーディオフレームの再生開始時刻はビデオピクチャの再生開始時刻よりも早いため、このようなときには処理はステップ S 4 0 9 に進む。

ステップ S 4 0 9 では、1 フレーム分のオーディオデータが破棄される。破棄する理由は、第 1 区間の映像と第 2 区間の映像とを連続して（途切れることなく）再生するため、第 2 区間のビデオピクチャ再生開始前のオーディオフレームは不要だからである。ステップ S 4 0 9 において 1 フレーム分のデータが破棄されると、再びステップ S 4 0 8 に戻って判定が行われる。これにより、図 1 1 の第 2 区間ににおいて、時刻 T c から時刻 T d までに再生されるオーディ

オフレームデータが破棄される。その後、不要なオーディオデータが破棄されると、処理はステップS403に進み、AV同期処理（ステップS403）が行われた後、ビデオピクチャおよびオーディオフレームが同期して出力される（ステップS404）。

5 次に、ステップS410では、オーディオ識別情報n<sub>a</sub>の値がビデオ識別情報n<sub>v</sub>の値よりも大きいか否かが判定される。前提として、ステップS410の前に行われたステップS407ではオーディオ識別情報の値がビデオ識別情報の値と異なっていると判定されている。オーディオ識別情報の値がビデオ識別情報の値よりも大きい場合にはステップS411に進み、小さい場合にはステップS412に進む。

例えば図12に示すように、第2区間のビデオストリームに対応するオーディオストリームが欠落しているとき、オーディオ識別情報の値がビデオ識別情報の値よりも大きくなり得る。また、その逆に、第2区間ににおいてビデオストリームが欠落し、オーディオストリームのみが存在するときには、ビデオ識別情報の値がオーディオ識別情報の値よりも大きくなり得る。

ステップS410の判定の結果に応じて処理がステップS411に進むと、AV同期制御部118はオーディオフレームデータの出力を待機させ、ビデオピクチャデータのみを出力して、映像のみが再生されるように制御する。そして、ビデオ信号を1フレーム出力した後は、処理はステップS307（図17）に戻る。

なお、仮にステップS410の判定を行わないとすると、図12

5

に示すようなケースでは、ビデオ出力部 117 が第 2 区間のビデオストリーム 703 に対応するビデオピクチャを出力している間に、オーディオ出力部は第 3 区間のオーディオストリーム 704 の復号およびオーディオフレームデータの出力を行ってしまう。一方、識別情報の値を利用するとこのような誤った再生を防いでいる。

10

ステップ S412 では、AV 同期制御部 118 は、その区間の終端までのオーディオフレームデータを破棄する。映像を連続的に再生するには、第 2 区間の音声を出力する必要はないからである。第 2 区間のオーディオフレームデータを破棄した後は、処理はステップ S306（図 17）に戻る。

15

なお、図 12 の例において第 2 区間のビデオストリームの映像が再生されているときのその後の処理を説明すると、まず、ステップ S412 から戻ったステップ S306 では、オーディオ識別情報 n<sub>a</sub> は第 3 区間の TS に対して割り当てられたオーディオ識別情報になる。よって、その後の処理は、ステップ S401、ステップ S402 およびステップ S405 へと進む。ただし、ステップ S405 では、第 2 区間には破棄すべきオーディオデータが存在しないため、そのまま処理が行われたとして再びステップ S306（図 17）に戻る。これを繰り返し、ステップ S402 の判定の結果、ステップ S403 およびステップ S404 に進むことにより、ステップ S404 において、第 3 区間のオーディオフレームデータおよびビデオピクチャデータが同期して出力される。

20

なお、一般的に、図 12 における第 2 区間と第 3 区間のオーディ

5

オ信号およびビデオ信号の再生時刻は独立している。すなわち、一般的には境界前後の再生時刻情報の絶対値の間には、全く相関はない。よって、再生時刻情報の絶対値の値が近似しているために第2区間の始点と第3区間の先頭の再生指定時刻が、偶然にも非常に近いとき、ステップS407およびS410の判定が存在しなければ、AV同期制御部118が第2区間のビデオデータと第3区間のオーディオデータを同期させて出力するよう制御するという不具合が発生する可能性がある。よって、ステップS407およびS410の判定が重要である。

10

15

さらに、取得したオーディオストリームに関連付けられた識別情報が、すでに出力されてしまった映像のビデオストリームに関連付けられた識別情報と一致するときは、AV同期制御部118はそのオーディオストリームに基づく音声の再生を停止させてもよい。そして、現在出力の対象としているビデオストリームに関連付けられた識別情報と同じ識別情報を有するオーディオストリームに基づいて音声を出力してもよい。

20

上記の説明では、AV同期制御部118が、オーディオ出力部110およびビデオ出力部117を制御するとして説明した。しかし、AV同期制御部118がオーディオ復号部108およびビデオ復号部115を制御するとしてもよい。そのとき、オーディオフレームデータの破棄に代えて、オーディオストリームを破棄することになる。その結果、復号を行う必要がなくなるので、再生装置の演算処理量を削減できる。ただし、出力されるオーディオフレームデータ

およびビデオピクチャデータを正確に同期させるためには、オーディオ出力部 110 およびビデオ出力部 117 に対する出力前のAV 同期処理は必要である。なお、データの破棄は、再生時刻情報に基づいてストリーム解析部 103 が行ってもよい。

5 また、本実施形態では、ダミーパケット挿入部 102 が識別情報を更新するとして説明した。しかし、識別情報の更新は、ストリーム解析部 103 が行ってもよい。ストリーム解析部 103 が識別情報を更新するとして再生装置 100 を構成すると、ダミーパケット挿入部 102 は、“continuity\_counter” フィールド 92 に固定値を記述したダミーパケットを挿入すれば十分になる。そしてストリーム解析部 103 がダミーパケットを検出する都度、内部的に識別情報の値を更新し、オーディオおよびビデオストリームの各区間のアドレスと識別情報を関連付ければよい。

10 また、これまで第 1 区間と第 2 区間の間にダミーパケットを挿入するとして説明したが、ダミーパケットを第 1 区間の前に挿入してもよい。これにより、第 1 区間のオーディオストリームまたはビデオストリームが欠落している場合に、識別情報が異なるオーディオフレームとビデオピクチャが同時に再生されるという不具合を回避できる。

15 ここで、図 19 および図 20 を参照しながら、ダミーパケットを第 1 区間の前に挿入することによって実現できる新たな処理を説明する。この処理は、第 1 区間終端のオーディオフレームの再生指定時刻を利用して、境界手前から境界に至るまでの音声を適切なタイ

ミングでフェードアウトするための処理である。また、第2区間先頭の音声をフェードインすることも可能である。

図19は、音声を区間終端においてフェードアウト再生し、区間始端においてフェードイン再生するための、ゲインとオーディオストリームとの関係を示す第1の例である。  
5

第1区間の音声が最後に再生される時刻Tcが、第1区間の映像が最後に再生される時刻Tdよりも早いとき、音声の再生が先に終了する。したがって、映像の再生時刻(Td)との関係に基づいて時刻Tbから音声のフェードアウトを開始すると、音声のフェードアウトが完了しないまま時刻Tcにおいて音声の再生が終了してしまう。その結果、異音が発生することがある。よって、波形Saのオーディオゲインは適切ではない。  
10

そこで、第1区間の音声が最後に再生される時刻Tcが、第1区間の映像が最後に再生される時刻Tdよりも早いときは、音声が最後に再生される時刻Tcとの関係で時刻Tbよりも早い時刻Taからフェードアウトを開始する。そのときのオーディオゲインは、波形Saに示すように、音声の再生が終了する時刻Tcにその出力ゲインも0になる。換言すれば、音声の再生が終了する時刻Tcに出力ゲインを0にできる時刻(Ta)から、フェードアウトを開始すればよい。時刻Taを決定する際の他のパラメータは、再生時のオーディオゲインの値およびフェードさせる時間である。  
15  
20

上述の処理を実現するために、図9(b)に示すように、第1区間の前に挿入する特定の符号に、オーディオ再生終端時刻の値を格

納する領域（audio\_end\_PTS）を定義し、音声の再生終了時刻  $T_c$  を再生時刻情報 PTS によって記述しておく。

A V 同期制御部 118 は、ビデオ再生終端時刻  $T_d$  よりもオーディオ再生終端時刻  $T_c$  が早いか否かを判定し、早いときは、時刻  $T_c$  からフェード時間だけ手前の時刻  $T_a$  からオーディオ出力部 110 に指示してフェードアウトを開始する。その結果、オーディオゲインの波形  $S_a$  は図 19 に示すように減少して時刻  $T_c$  において 0 になる。これにより、第 1 区間の終端のフェードアウト処理による異音発生を防ぐことができる。

また、境界後にオーディオ信号の再生をする際には、先に説明したように、第 2 区間の音声の再生開始時刻  $T_e$  が映像の再生開始時刻  $T_0 f$  よりも早いときには時刻  $T_e$  から時刻  $T_f$  までのオーディオストリームが破棄される。この場合、第 1 区間の映像と第 2 区間の映像とを途切れなく再生するためには、A V 同期制御部 118 は映像の再生開始時刻  $T_f$  から音声のフェードインを開始すればよい。

図 20 は、音声を区間終端においてフェードアウト再生し、区間始端においてフェードイン再生するための、ゲインとオーディオストリームとの関係を示す第 2 の例である。音声の再生終端時刻  $T_d$  が映像の再生終端時刻  $T_c$  よりも遅い場合には、上述のとおり、映像の再生終端時刻  $T_c$  後、時刻  $T_d$  までの音声は再生されない。したがって、第 1 区間の音声のフェードアウトを音声の再生終端時刻 ( $T_d$ ) との関係に基づいて開始すると、映像の再生終端時刻  $T_c$  までにフェードアウトが完了せず、境界で異音が発生する場合があ

る。よって、波形 S b のオーディオゲインは適切ではない。

そこで、図 9 (b) に示すように、ダミーパケットにビデオ／オーディオ再生終端時刻を格納する領域をそれぞれ設ける。AV 同期制御部 118 は、それらの領域からビデオ再生先頭時刻 T c とオーディオ再生終端時刻 T d の情報を読み出し、オーディオ再生終端時刻 T d の方が遅い場合には、映像の再生終端時刻 T c を基準としてフェード時間だけ早い時刻 T a からフェードアウトを開始する。これにより、境界前において映像の再生終了とともにフェードアウトが完了し、異音を発生することなく第 2 区間に映像および音声を切り替えることができる。

また、境界後の第 2 区間の音声の再生を開始する場合、区間の先頭で音声をフェードインすると、一般にはユーザは聴きやすくなる。ここで、境界後のオーディオの再生開始時刻 T f がビデオの再生開始時刻 T e よりも遅いときには、波形 S c のオーディオゲインで音声が出力されるよりも前からゲインを上げてフェードイン処理を開始すると、実際に音声の再生先頭時刻 T f において突然振幅の大きい信号が出力されることになる。すると、出力機器であるスピーカを破損する危険がある。このような事例は、例えば、AV 同期制御部 118 がオーディオ復号部 108 によるオーディオストリームの復号を時刻 T f まで待機させ、かつ、オーディオゲインの調整をオーディオ出力部 110 で行うように構成された再生装置 100 で起こる可能性がある。

そこで、図 9 (b) に示すように、ダミーパケットにビデオ／オ

ーディオ再生先頭時刻を格納する領域をそれぞれ設ける。AV同期制御部118は、それらの領域からオーディオ再生先頭時刻 $T_f$ とビデオ再生先頭時刻 $T_e$ の情報を読み出し、オーディオ再生先頭時刻 $T_f$ が大きい場合にはオーディオ信号のフェードインを時刻 $T_f$ から開始するようすればよい。

以上のように、システムストリームの先頭に挿入するダミーパケットに、オーディオおよびビデオの再生先端時刻および再生終端時刻をそれぞれ格納することにより、境界近傍における異音を発生すことなく、フェードインおよびフェードアウトを行うことが可能となる。

以上、本発明のデータ処理装置の実施形態である、再生装置100の構成および動作を説明した。なお、図7において、ストリーム読み出し部101はスイッチとして表わされているが、これは機械的あるいは電気的なスイッチに限るものではない。ストリーム読み出し部101は、記録媒体の任意のアドレスを指定してTSを読み出すことができればよい。

なお、再生装置100の各入力バッファ105、106、112および113の構成は従来例と同様の構成で示しているが、第1のオーディオ入力バッファ105と第2のオーディオ入力バッファ106とは一つのバッファであってもよい。同様に、第1および第2のビデオ入力バッファ112、113も一つのバッファであってもよい。その場合、第1区間のTSから抽出されたオーディオストリームに続けて第2区間のTSから抽出されたオーディオストリーム

を格納し、上で説明したように、例えば図16に示すようなテーブルによって、各区間のオーディオストリームを格納したアドレスとストリーム解析部103によって各区間のオーディオストリームに割り当てられた識別情報nと再生指定時刻とを読み出すことができればよい。

なお、上記の説明では、ダミーパケット71はTSパケットのひとつとして説明したが、これに限るものではない。ストリーム解析部103がオーディオデータおよびビデオデータと識別可能なデータであればよい。また光ディスク120に代えて、ハードディスク(図示せず)を利用することができる。光ディスク120には必ず複数のTSが記録されていなくてもよく、例えば1つのTSの異なる区間を、それぞれ別個のTSとして捉え、上述の処理を適用することもできる。

## 15 産業上の利用可能性

本発明によれば、複数のデータストリームを連続的に再生する際に、音声と映像とをずれることなく同期して再生することができるデータ処理装置が提供される。

## 請 求 の 範 囲

1. 映像データおよび音声データを含むデータストリームから映像および音声を再生するデータ処理装置であって、前記映像データおよび前記音声データの各々には再生のタイミングを示す時刻情報が付加されており、

第1データストリームおよび第2データストリームを連続して取得するストリーム取得部と、

前記第1データストリームと前記第2データストリームとが切り替わるデータ位置に、境界を示すダミーデータを挿入する挿入部と、  
前記ダミーデータを検出して、前記第1データストリームおよび前記第2データストリームに異なる識別情報を割り当て、各データストリームの前記映像データおよび前記音声データに前記識別情報を関連付ける解析部と、

同じ識別情報が関連付けられた映像データおよび音声データの各時刻情報に基づいて、前記映像データに基づく映像および前記音声データに基づく音声の出力タイミングを制御する制御部と、

前記出力タイミングで、前記映像および前記音声を出力する出力部と

20 を備えたデータ処理装置。

2. 前記制御部は、前記映像データに付加された時刻情報および前記音声データに付加された時刻情報に基づいて、前記第1データ

ストリームに関する前記映像および前記音声の各再生終了時刻を特定し、前記音声の再生終了時刻が前記映像の再生終了時刻よりも遅い時刻であるときは、前記映像の再生終了時刻から前記音声の再生終了時刻までの前記音声の出力を停止する、請求項 1 に記載のデータ処理装置。

3. 前記制御部は、前記映像データに付加された時刻情報および前記音声データに付加された時刻情報に基づいて、前記第 2 データストリームに関する前記映像および前記音声の各再生開始時刻を特定し、前記音声の再生開始時刻が前記映像の再生開始時刻よりも早い時刻であるときは、前記音声の再生開始時刻から前記映像の再生開始時刻までの前記音声の出力を停止する、請求項 1 に記載のデータ処理装置。

15 4. 前記制御部は、映像データおよび音声データに関連付けられた識別情報が異なるとき、前記映像データに基づく映像のみを先に出力させ、前記映像の再生終了後に得られた映像データに関連付けられた識別情報が前記音声データに関連付けられた識別情報と一致したとき、同じ識別情報が関連付けられた前記映像データおよび前記音声データの各時刻情報に基づいて、前記映像データに基づく映像および前記音声データに基づく音声の出力タイミングを制御する、請求項 1 に記載のデータ処理装置。

5. 前記ストリーム取得部は、3以上のデータストリームを連続して取得することが可能であり、

前記挿入部は、連続して取得される2つのデータストリームが切り替わるデータ位置ごとに、識別情報に対応する、単調増加または  
5 単調減少させた値を有するダミーデータを挿入する、請求項1に記載のデータ処理装置。

6. 前記制御部は、前記音声データに関連付けられた識別情報が、過去に出力された映像の映像データに関連付けられた識別情報と一致するときは、前記音声データに基づく音声の出力を停止し、現在  
10 出力の対象としている映像の映像データに関連付けられた識別情報と同じ識別情報を有する音声データに基づいて音声を出力する、請求項5に記載のデータ処理装置。

15 7. 前記制御部は、前記音声データに関連付けられた識別情報がまだ出力されていない映像の映像データに関連付けられた識別情報と一致するときは前記音声データに基づく音声の出力を停止し、その後取得した映像データに関連付けられた識別情報と一致したときにおいて、前記音声データに基づいて音声を出力する、請求項1に記載のデータ処理装置。

20 8. 前記データストリームは、映像データを格納したパケットおよび音声データを格納したパケットを含むパケット構造で構成され

ており、

前記挿入部は、前記第1データストリームの最終パケットと、前記第2データストリームの先頭パケットとの間の位置に、境界を示すダミーパケットを挿入する、請求項1に記載のデータ処理装置。

5

9. 映像データおよび音声データを含むデータストリームから映像および音声を再生するデータ処理方法であって、前記映像データおよび前記音声データの各々には再生のタイミングを示す時刻情報が付加されており、

10 第1データストリームおよび第2データストリームを連続して取得するステップと、

前記第1データストリームと前記第2データストリームとが切り替わるデータ位置に、境界を示すダミーデータを挿入するステップと、

15 前記ダミーデータを検出して、前記第1データストリームおよび前記第2データストリームに異なる識別情報を割り当て、各データストリームの前記映像データおよび前記音声データに前記識別情報を関連付けるステップと、

20 同じ識別情報が関連付けられた映像データおよび音声データの各時刻情報に基づいて、前記映像データに基づく映像および前記音声データに基づく音声の出力タイミングを制御するステップと、

前記出力タイミングで、前記映像および前記音声を出力するステップと

を包含するデータ処理方法。

10. 前記制御するステップは、

前記映像データに付加された時刻情報および前記音声データに付  
5 加された時刻情報に基づいて、前記第1データストリームに関する  
前記映像および前記音声の各再生終了時刻を特定するステップと、

前記音声の再生終了時刻が前記映像の再生終了時刻よりも遅い時  
刻であるときは、前記映像の再生終了時刻から前記音声の再生終了  
時刻までの前記音声の出力を停止するステップとを包含する、請求  
10 項9に記載のデータ処理方法。

11. 前記制御するステップは、

前記映像データに付加された時刻情報および前記音声データに付  
15 加された時刻情報に基づいて、前記第2データストリームに関する  
前記映像および前記音声の各再生開始時刻を特定するステップと、

前記音声の再生開始時刻が前記映像の再生開始時刻よりも早い時  
刻であるときは、前記音声の再生開始時刻から前記映像の再生開始  
時刻までの前記音声の出力を停止するステップとを包含する、請求  
20 項9に記載のデータ処理方法。

12. 前記制御するステップは、

映像データおよび音声データに関連付けられた識別情報が異なる  
とき、前記映像データに基づく映像のみを先に出力させるステップ

と、

前記映像の再生終了後に得られた映像データに関連付けられた識別情報が前記音声データに関連付けられた識別情報と一致したとき、同じ識別情報が関連付けられた前記映像データおよび前記音声データの各時刻情報に基づいて、前記映像データに基づく映像および前記音声データに基づく音声の出力タイミングを制御するステップとを包含する、請求項 9 に記載のデータ処理方法。

1 3 . 前記ストリームを取得するステップは、3 以上のデータストリームを連続して取得し、

前記挿入するステップは、連続して取得される 2 つのデータストリームが切り替わるデータ位置ごとに、識別情報に対応する、単調増加または単調減少させた値を有するダミーデータを挿入する、請求項 9 に記載のデータ処理方法。

15

1 4 . 前記制御するステップは、

前記音声データに関連付けられた識別情報が、過去に出力された映像の映像データに関連付けられた識別情報と一致するときは、前記音声データに基づく音声の出力を停止するステップと、

20

現在出力の対象としている映像の映像データに関連付けられた識別情報と同じ識別情報を有する音声データに基づいて音声を出力するステップとを包含する、請求項 1 3 に記載のデータ処理方法。

15. 前記制御するステップは、

前記音声データに関連付けられた識別情報がまだ出力されていない映像の映像データに関連付けられた識別情報と一致するときは前記音声データに基づく音声の出力を停止するステップと、

5 その後取得した映像データに関連付けられた識別情報と一致したときにおいて、前記音声データに基づいて音声を出力するステップと

を包含する、請求項 9 に記載のデータ処理方法。

10 16. 前記データストリームは、映像データを格納したパケットおよび音声データを格納したパケットを含むパケット構造で構成されており、

前記挿入するステップは、前記第 1 データストリームの最終パケットと、前記第 2 データストリームの先頭パケットとの間の位置に、  
15 境界を示すダミーパケットを挿入する、請求項 9 に記載のデータ処理方法。

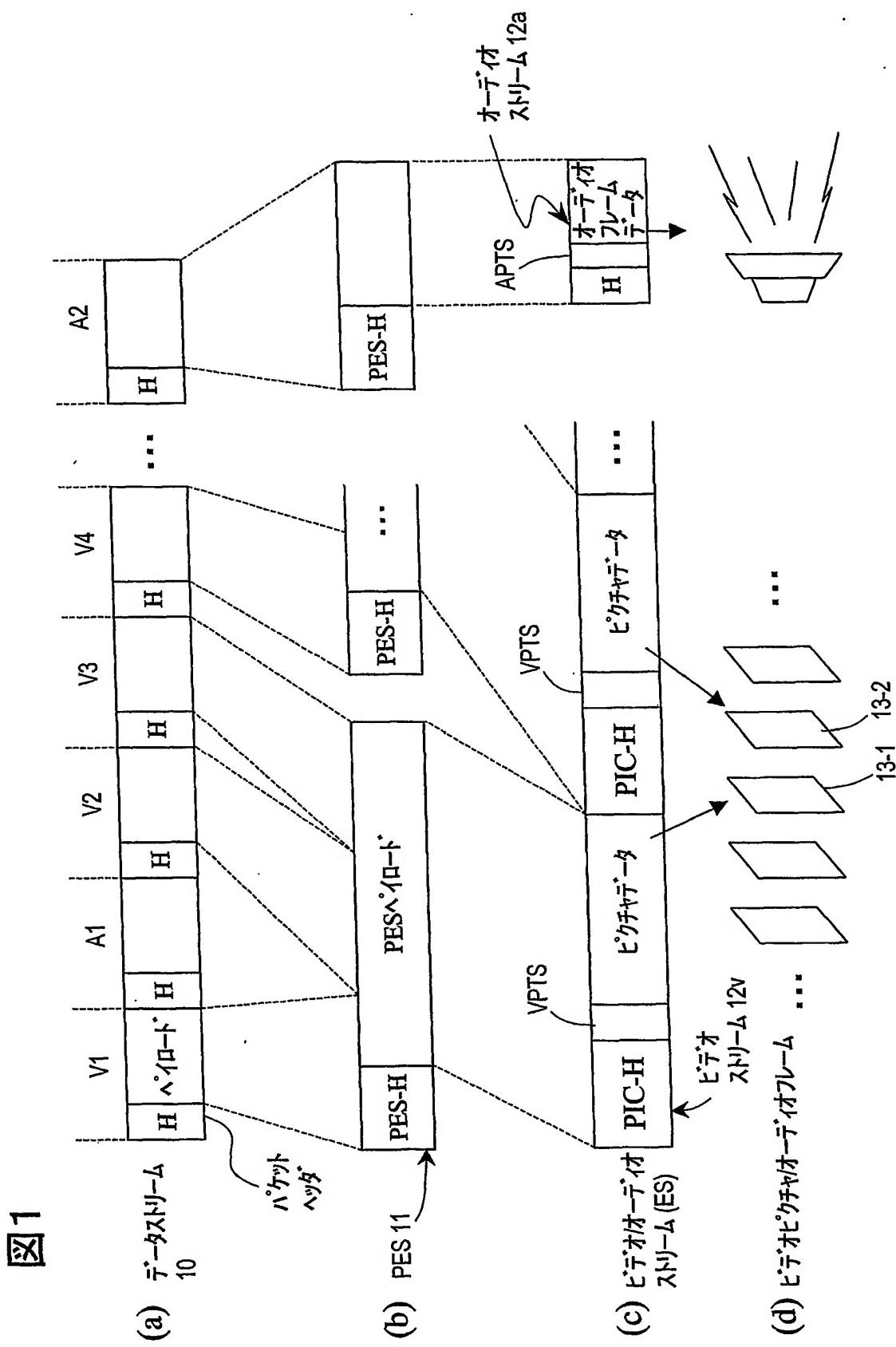


図2

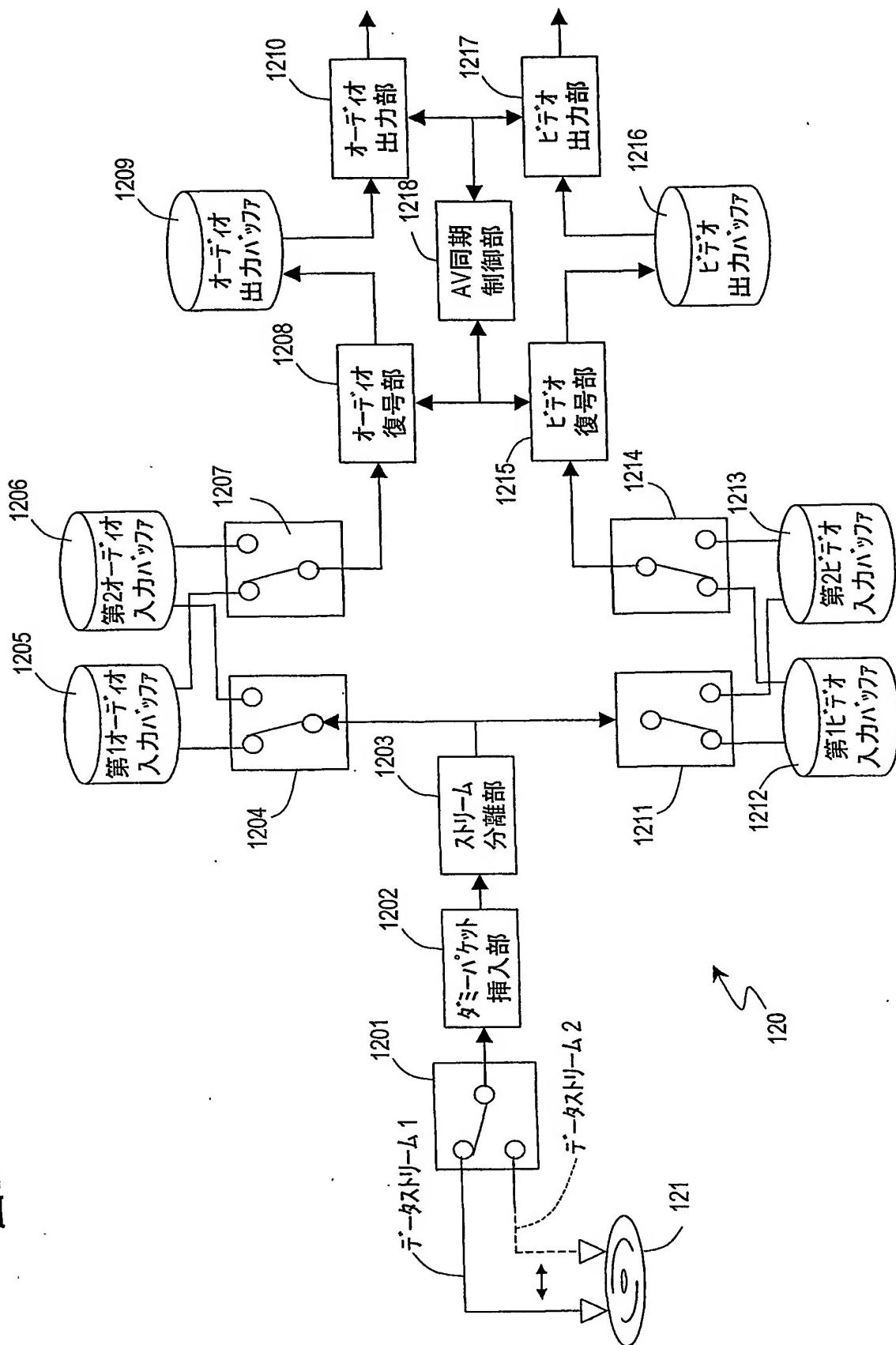


図3

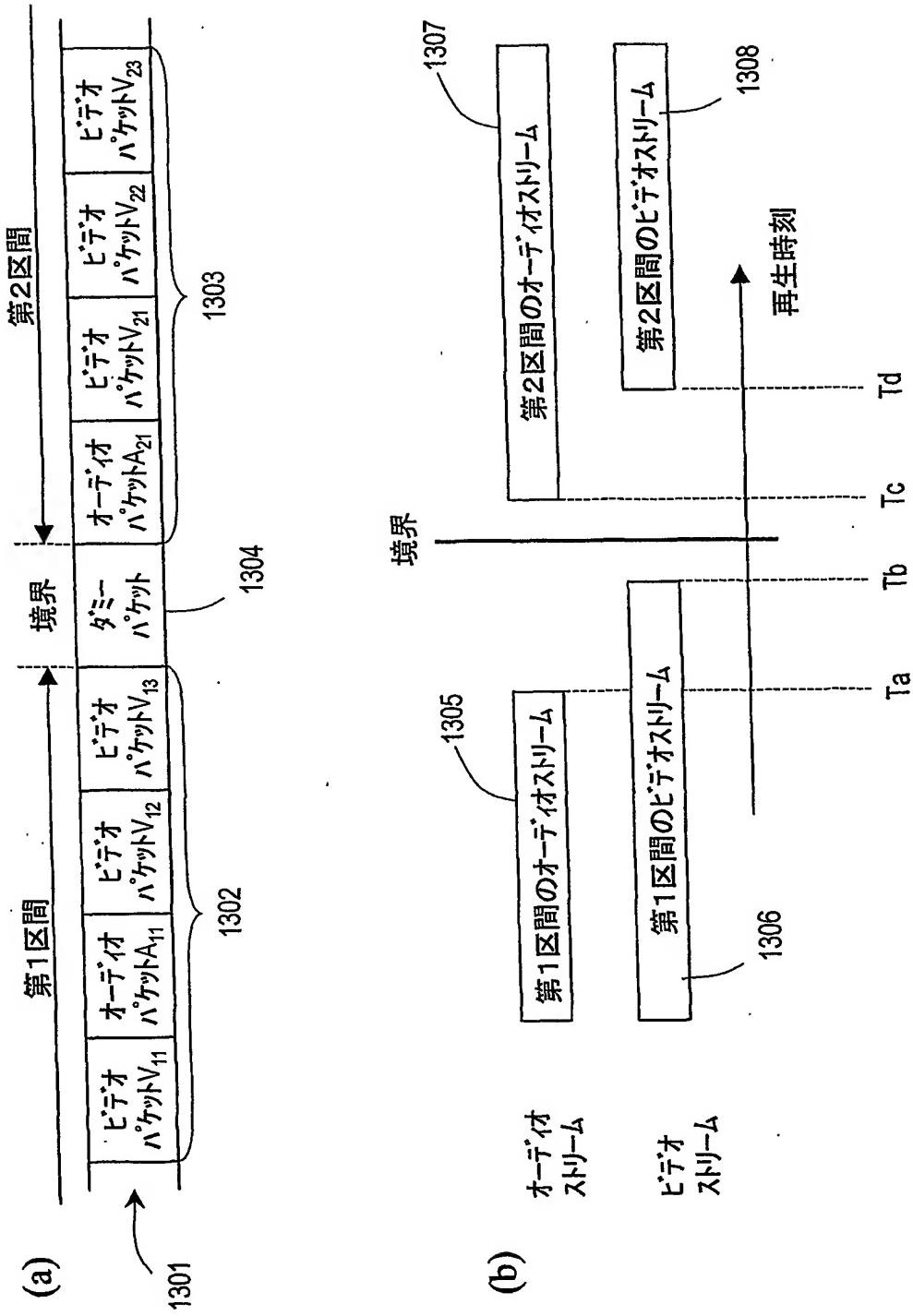


図4

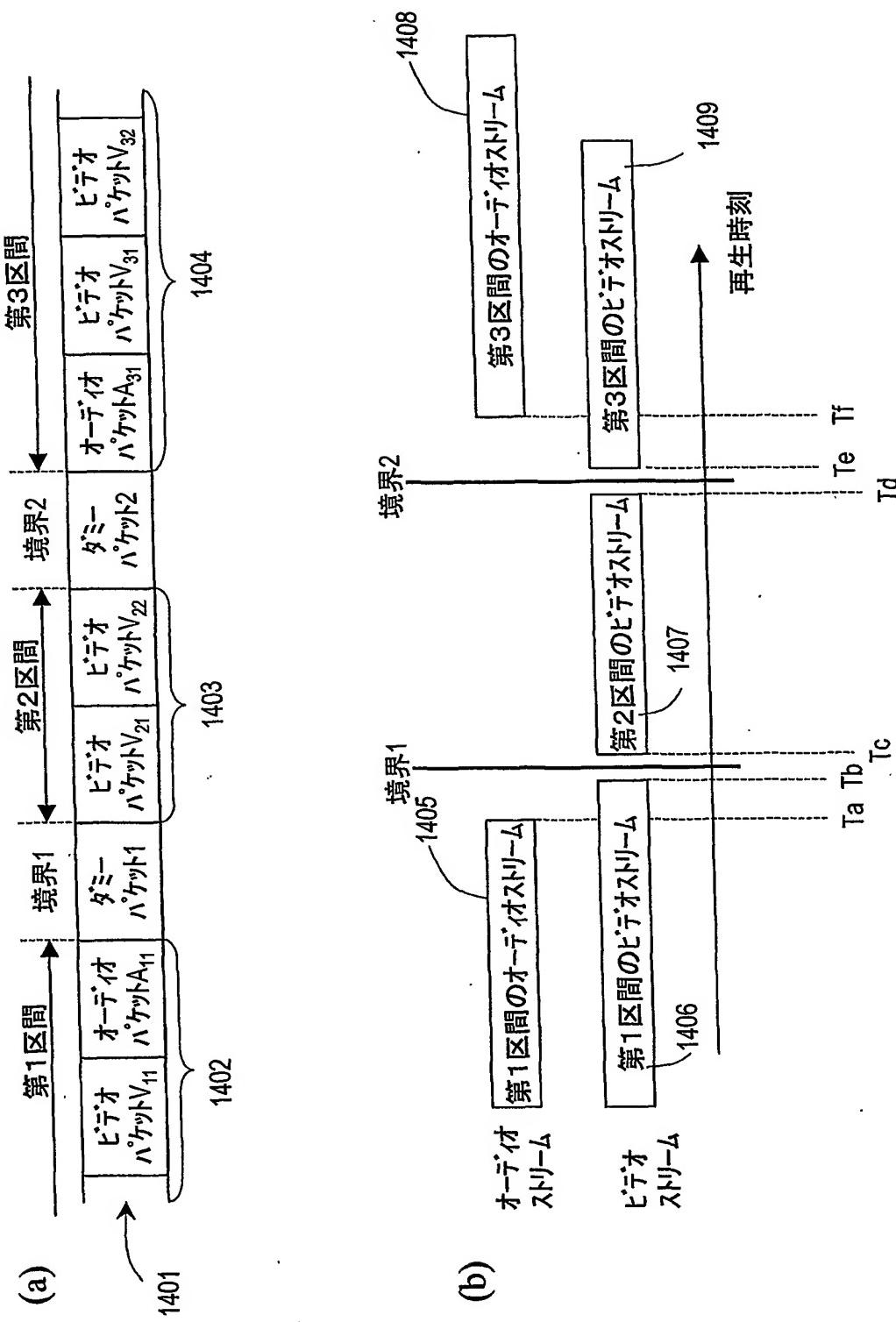
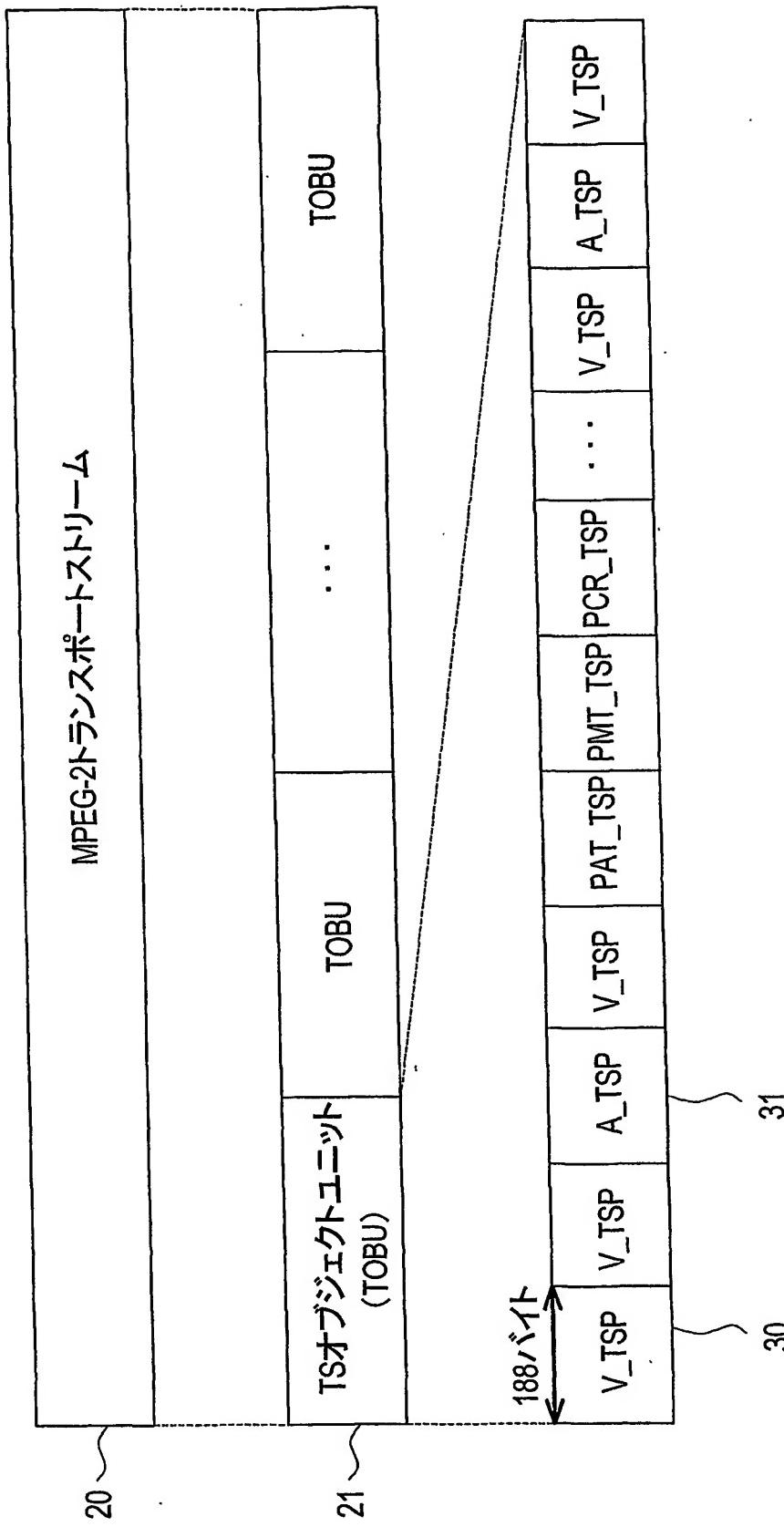
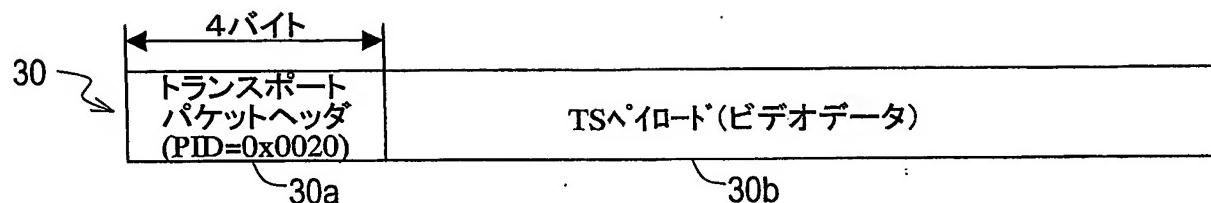


図5

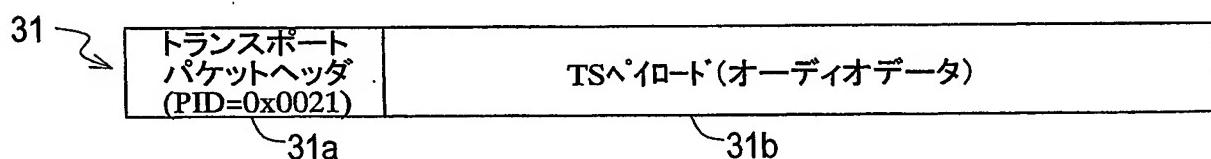


**図6**

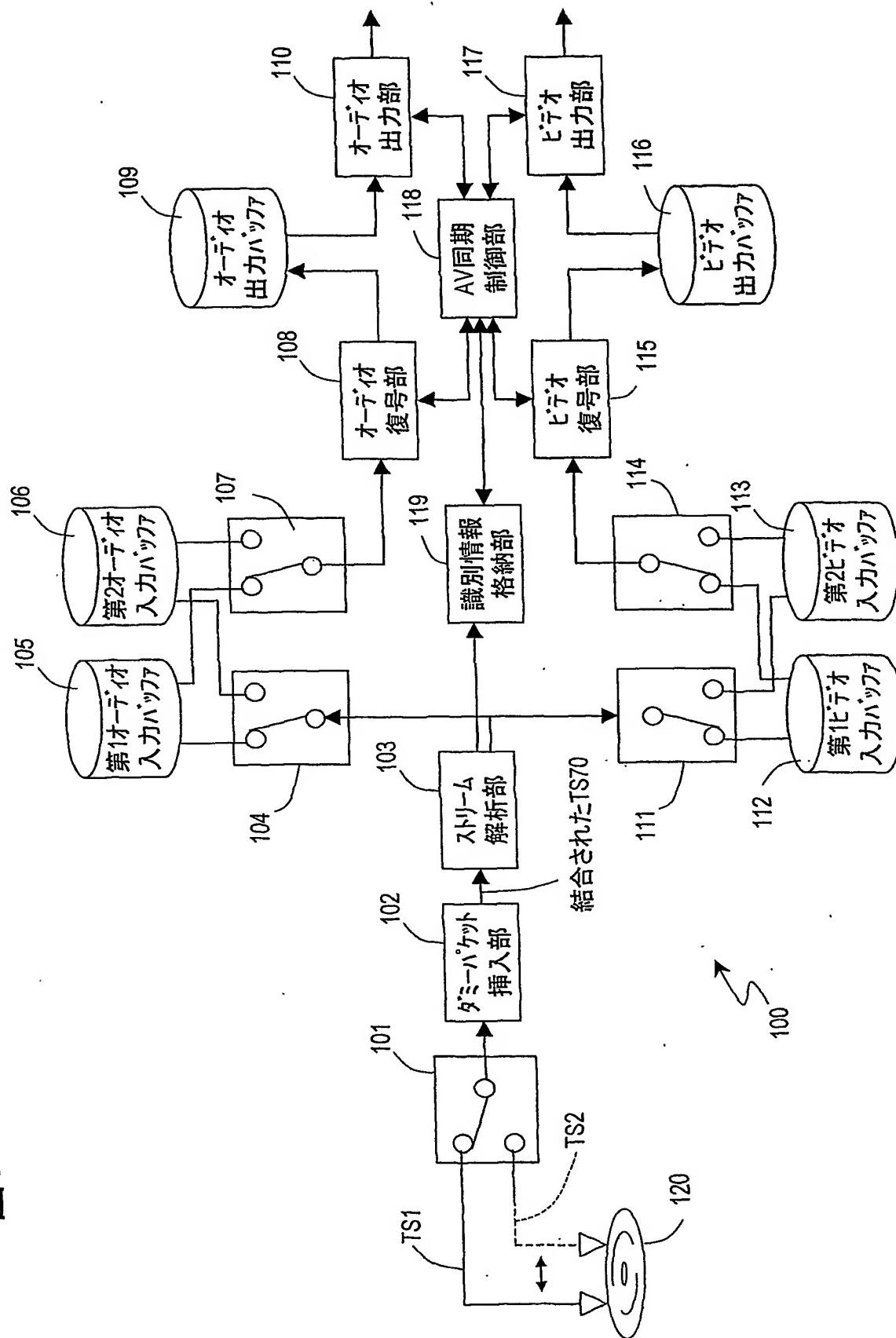
(a) V\_TSP



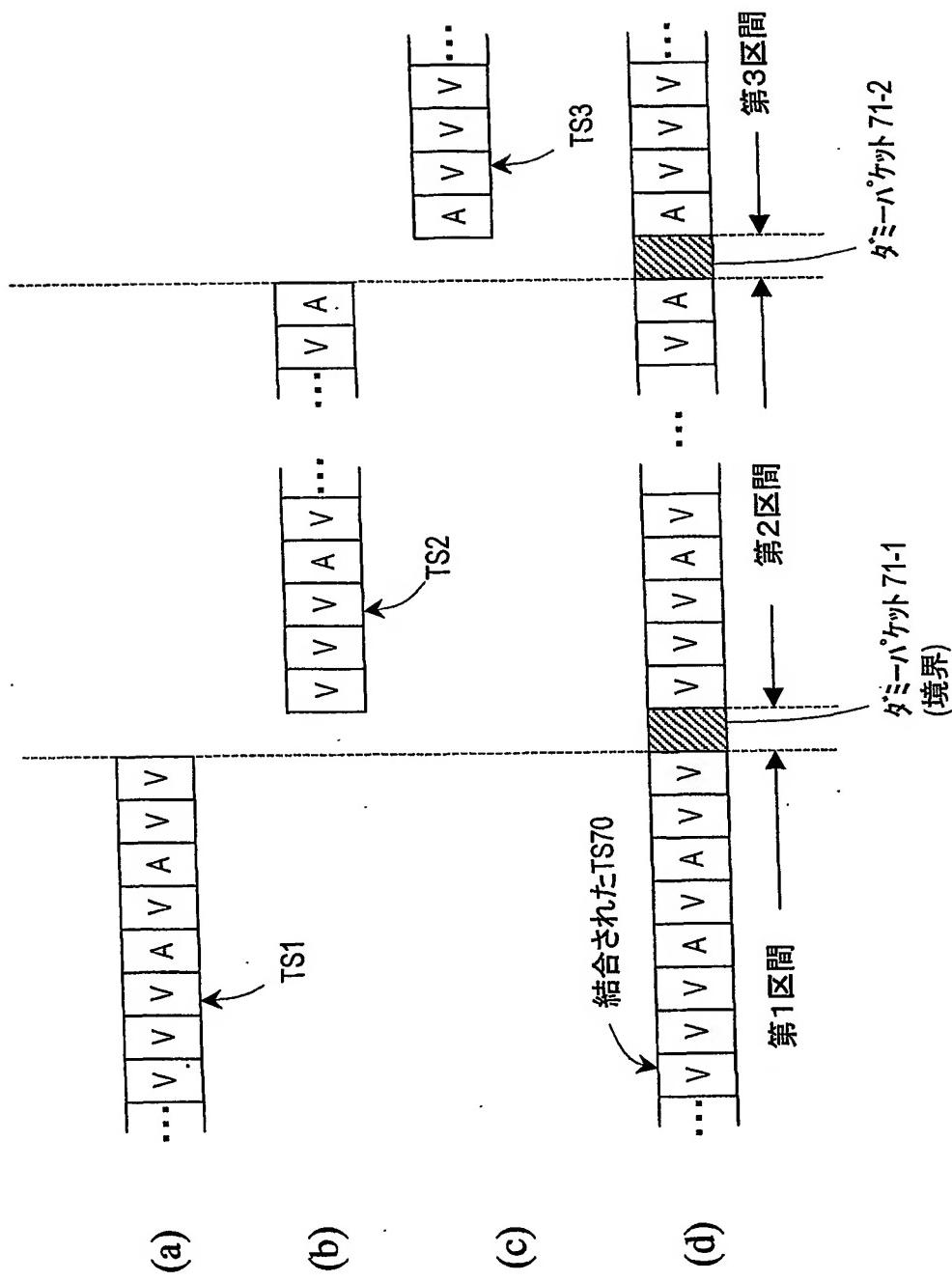
(b) A\_TSP



7

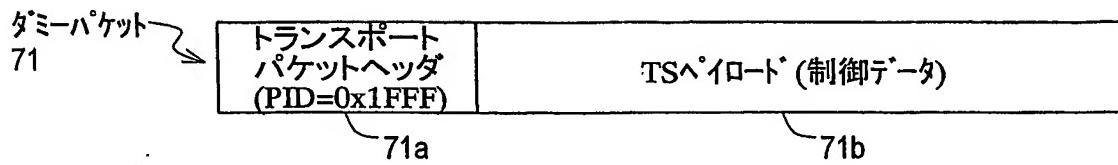


88



## 図9

(a)



(b)

	データ構造	ビット数	データ
トランSPORT パケットヘッダ	sync_byte	8	0x47
	transport_error_indicator	1	0b
	payload_unit_start_indicator	1	1b
	transport_priority	1	1b
	PID	13	0x1FFF
	transport_scrambling_control	2	00b
	adaptation_field_control	2	01b
	continuity_counter	4	識別情報(0x0~0xF)
TSペイロード	Dummy 判別情報	24	"DUM"(0x44 55 4D)
	Dummy_ID	4	0xF
	Reserved	12	0xFFF
	audio_start PTS	33	オーディオ再生先頭時刻
	video_start PTS	33	ビデオ再生先頭時刻
	audio_end PTS	33	オーディオ再生終端時刻
	video_end PTS	33	ビデオ再生終端時刻
	Stuffing_data	1300	0xFF, 0xFF, ...

図10

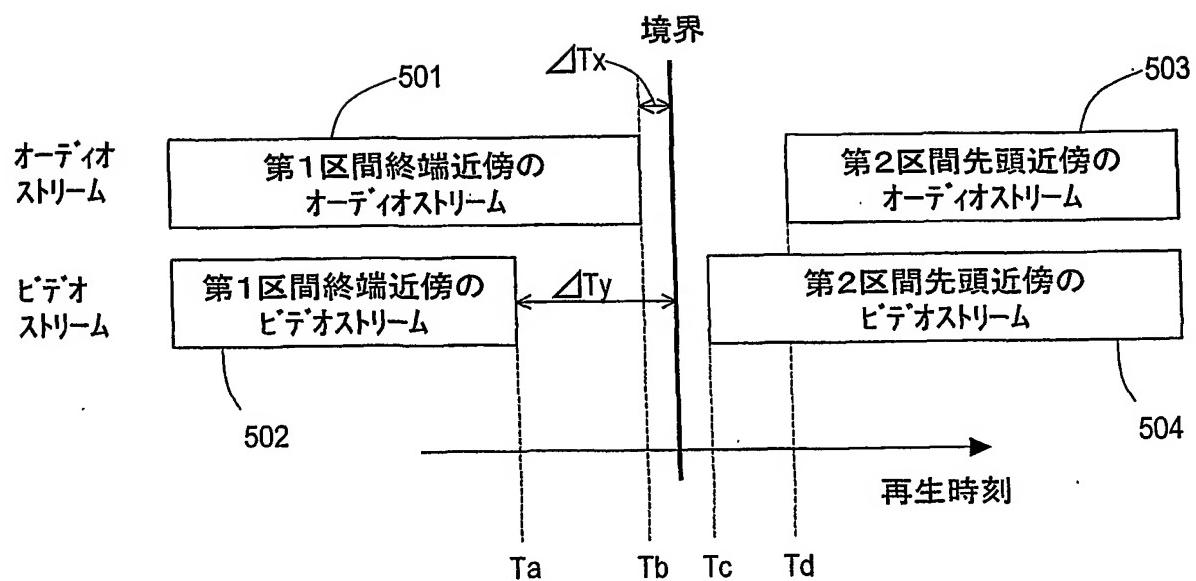


図11

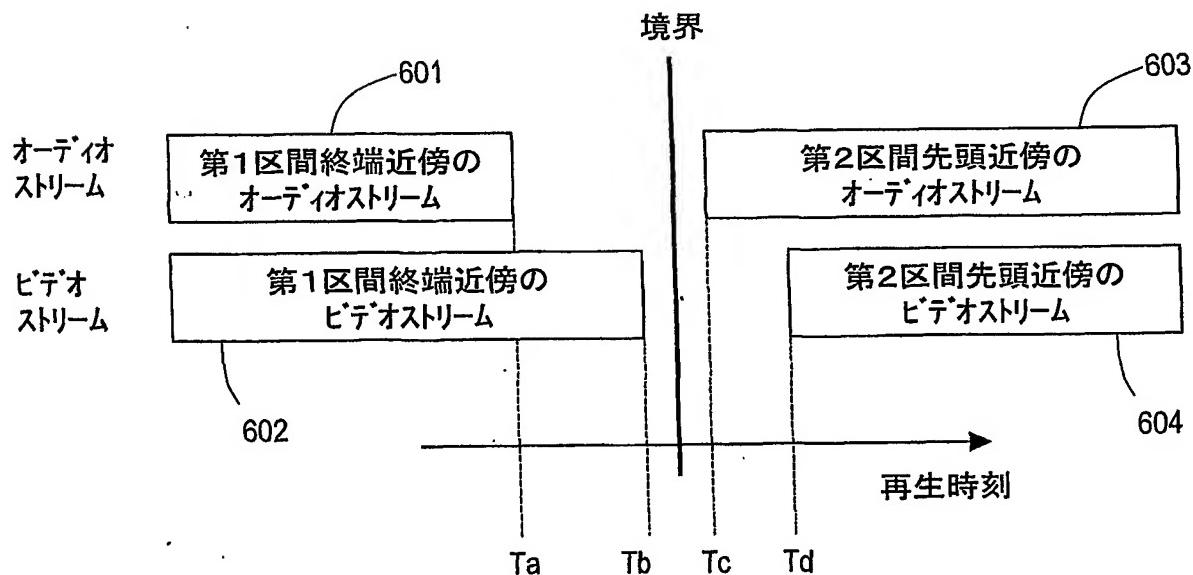


図12

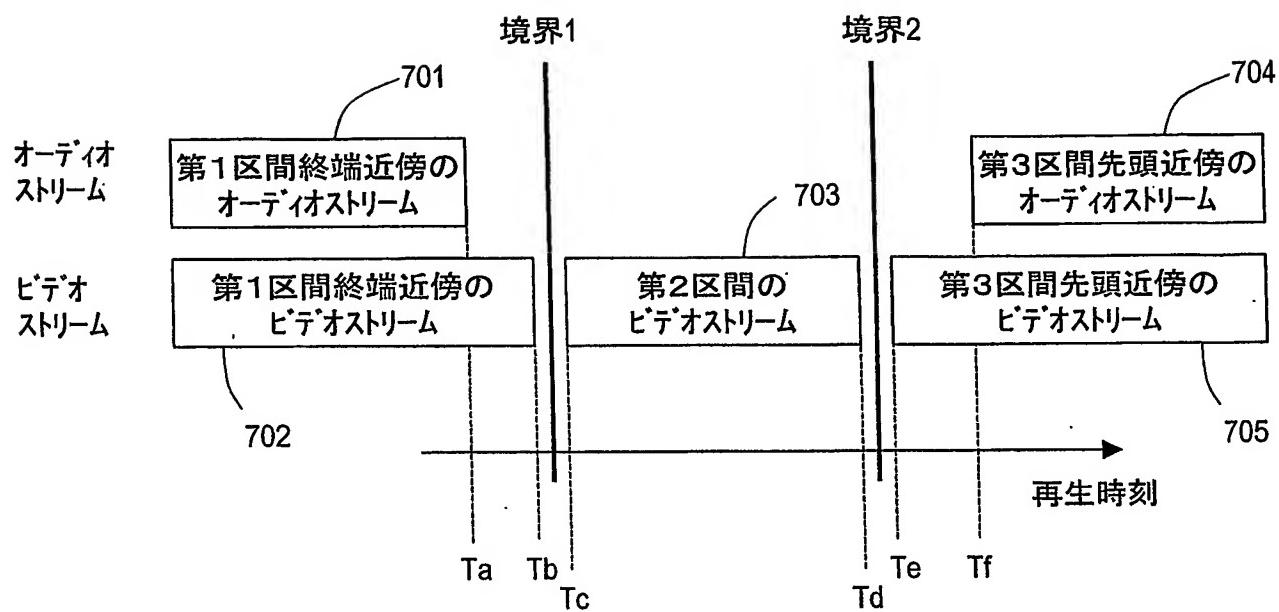


図13

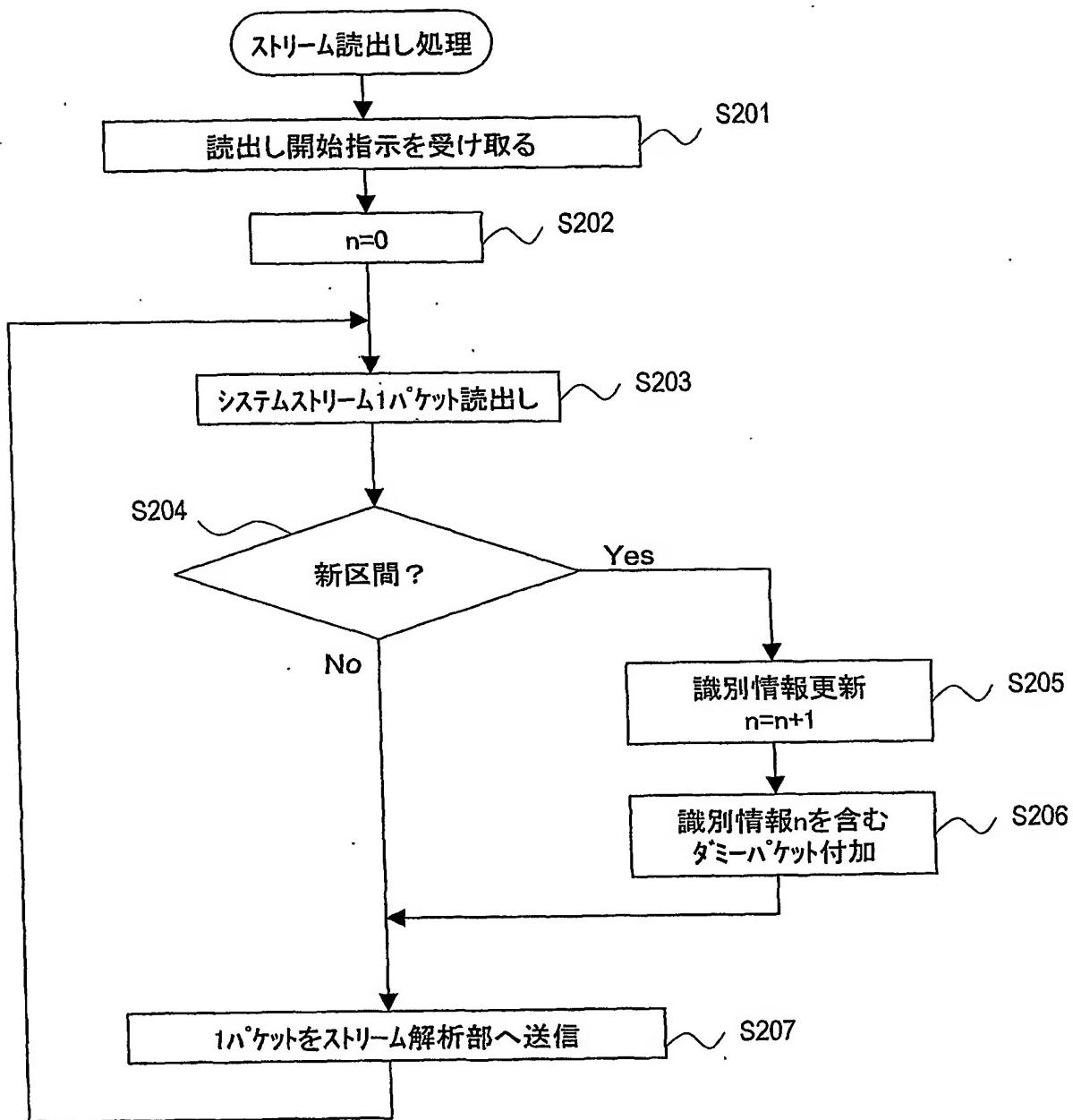


図14

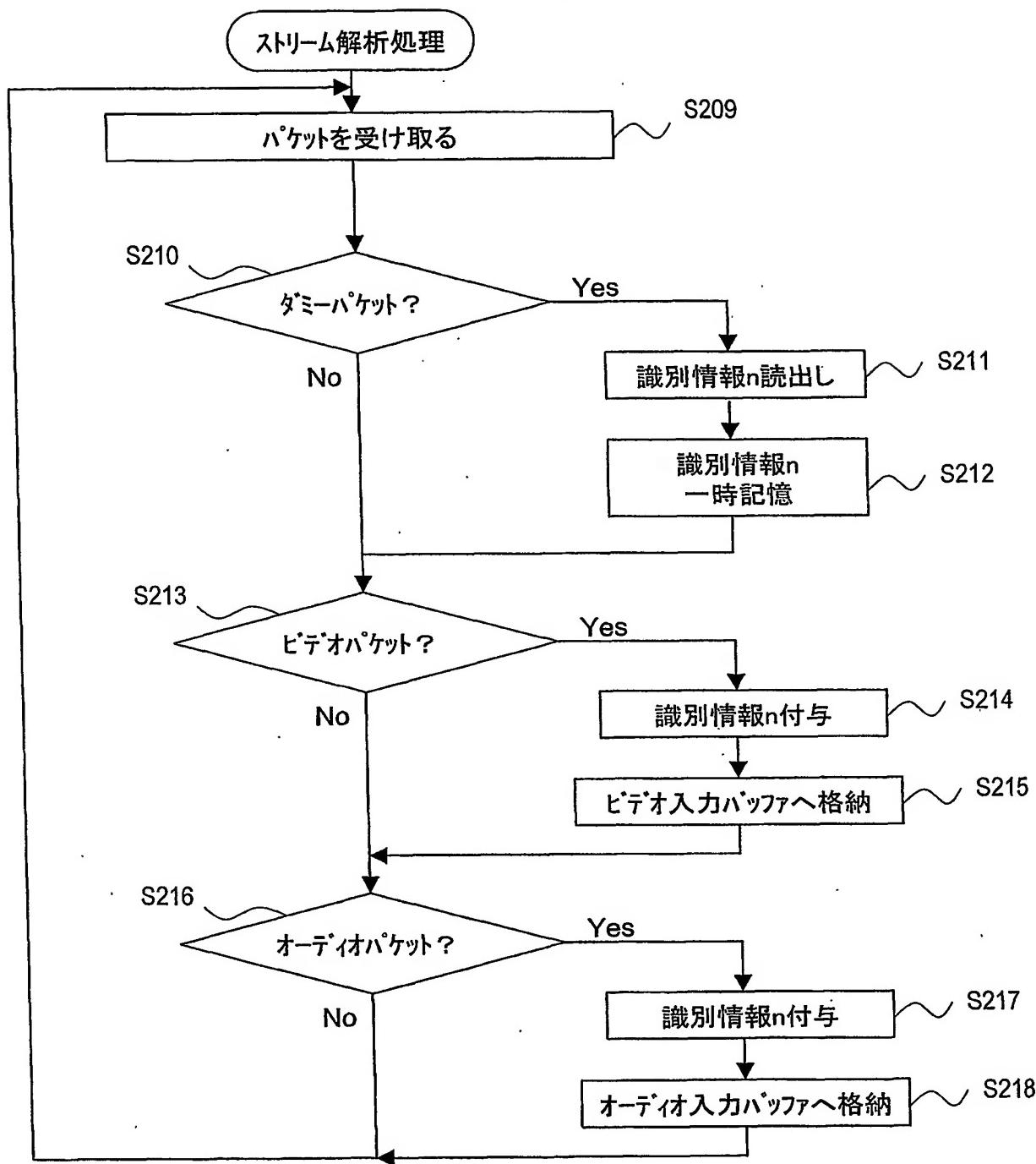


図15

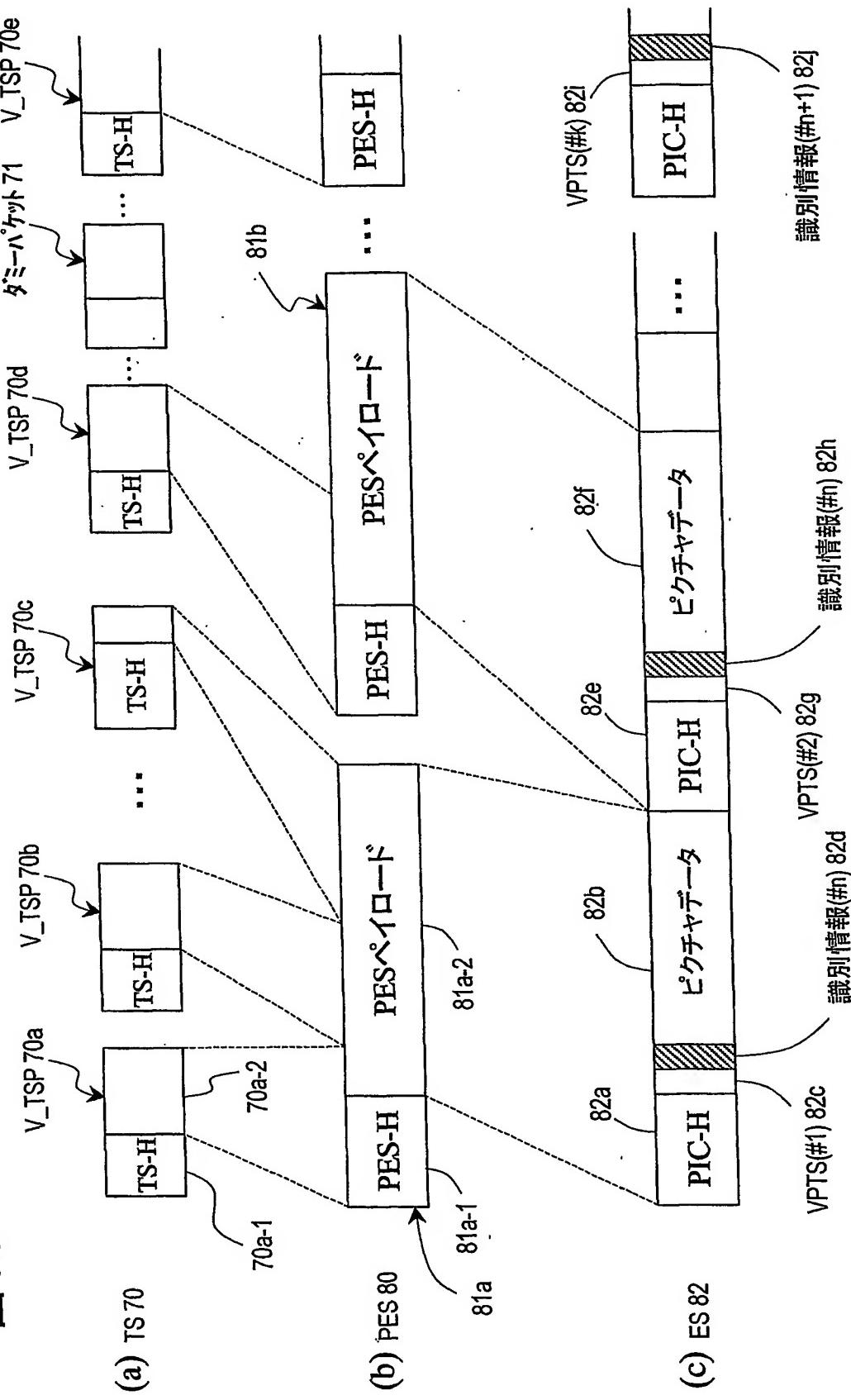


図16

識別情報	APTS	アドレス
1	APTS1	アドレス1
1	APTS2	アドレス2
2	APTS3	アドレス3
2	APTS4	アドレス4
2	APTS5	アドレス5
3	APTS6	アドレス6
3	APTS7	アドレス7
3	⋮	⋮
3	⋮	⋮

図17

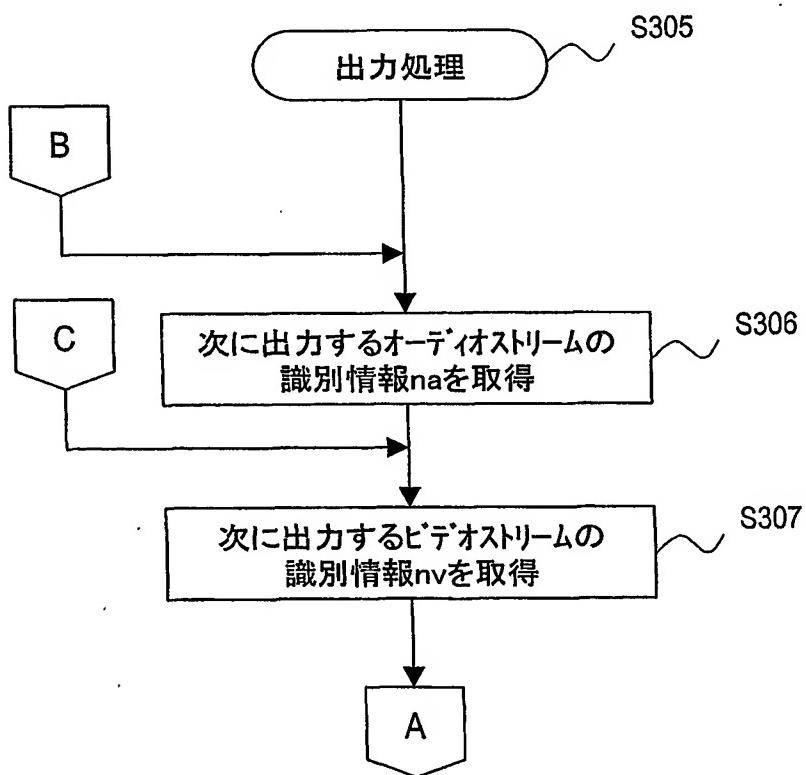


図18

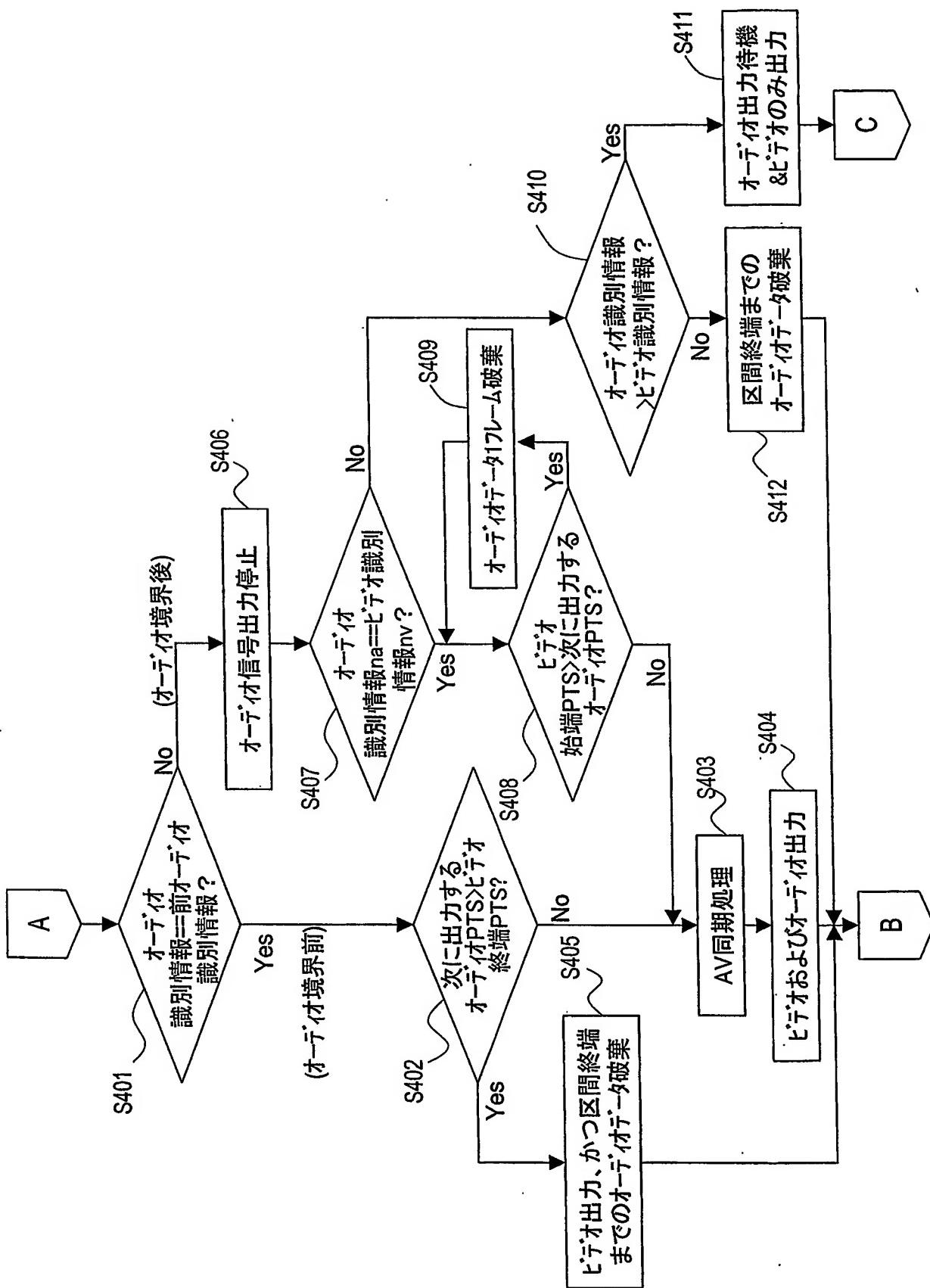


図19

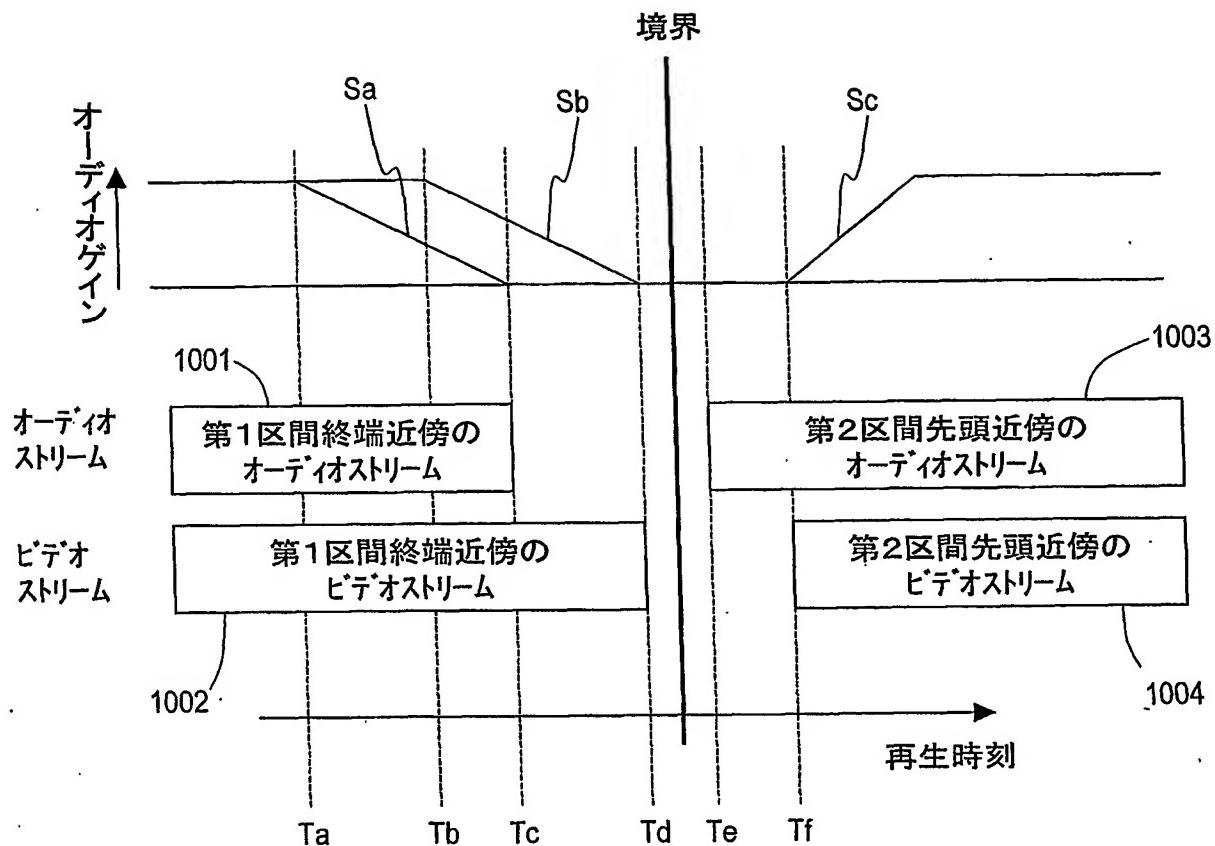
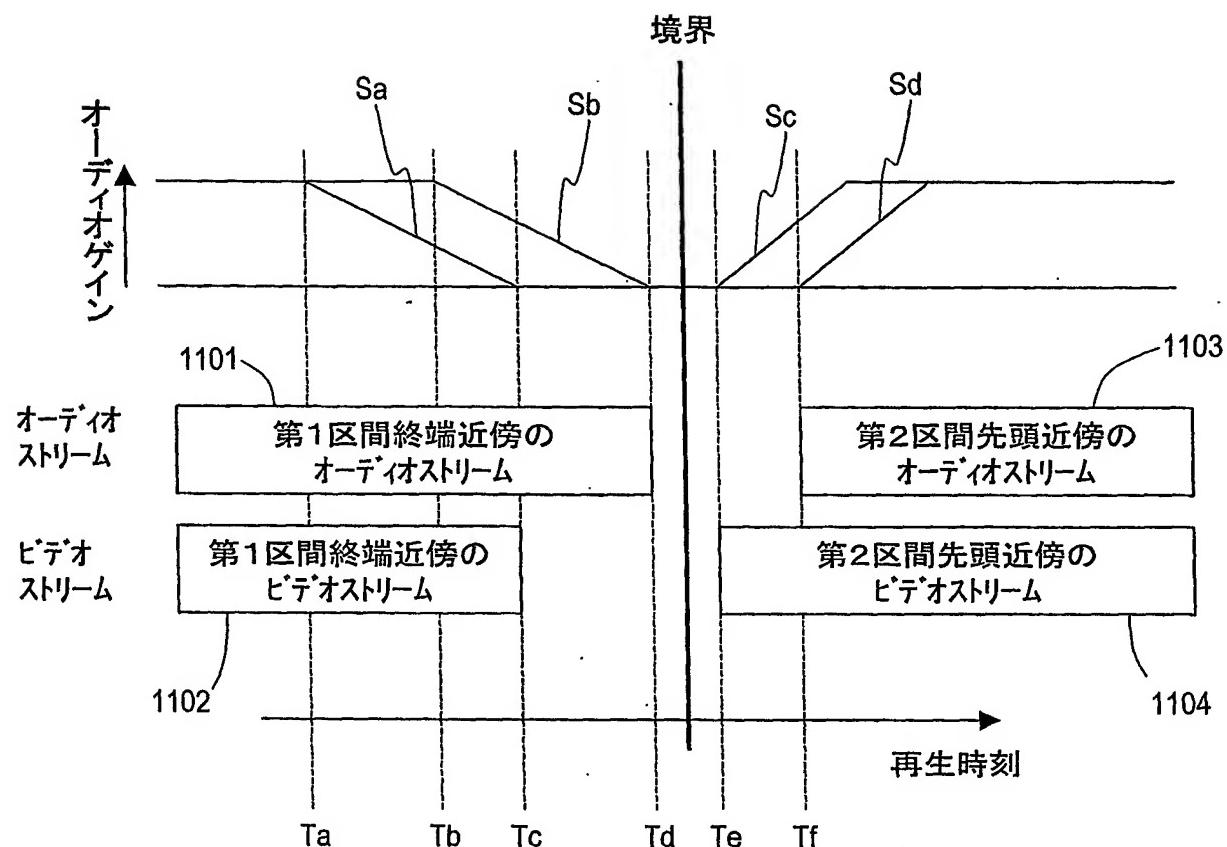


図20



## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2004/009522

## A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl<sup>7</sup> H04N5/92, H04N5/93, G11B20/10, G11B27/10

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl<sup>7</sup> H04N5/92, H04N5/93, G11B20/10, G11B27/10

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Toroku Jitsuyo Shinan Koho	1994-2004
Kokai Jitsuyo Shinan Koho	1971-2004	Jitsuyo Shinan Toroku Koho	1996-2004

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2003-87744 A (Matsushita Electric Industrial Co., Ltd.), 20 March, 2003 (20.03.03), Full text; all drawings & WO 97/13366 A1 & EP 847200 A1 & US 5884004 A1	1-5, 8, 9-13, 16
A	JP 2002-16880 A (Sony Corp.), 18 January, 2002 (18.01.02), Full text; all drawings (Family: none)	1, 8, 9, 16
A	JP 2002-199336 A (Toshiba Corp.), 12 July, 2002 (12.07.02), Par. No. [0338] (Family: none)	2-4, 10-12

 Further documents are listed in the continuation of Box C. See patent family annex.

\* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier application or patent but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search  
17 September, 2004 (17.09.04)Date of mailing of the international search report  
05 October, 2004 (05.10.04)Name and mailing address of the ISA/  
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

**INTERNATIONAL SEARCH REPORT**

International application No. PCT/JP2004/009522
--

Continuation of Box No.II-2 of continuation of first sheet(2)

Moreover, the audio and the video whose outputs are stopped are separate streams. It is unclear why correlated identification information is given to them.

**INTERNATIONAL SEARCH REPORT**

International application No.

PCT/JP2004/009522

**Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2.  Claims Nos.: 6, 7, 14, 15  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:  
The difference between "audio (data) whose output is stopped" having identification information correlated with video data outputted in the past (not yet outputted) and "audio (data) which has been outputted" is unclear.  
(Continued to extra sheet.)
3.  Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

1.  As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2.  As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3.  As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  
4.  No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

**Remark on Protest**

- The additional search fees were accompanied by the applicant's protest.  
 No protest accompanied the payment of additional search fees.

## A. 発明の属する分野の分類(国際特許分類(IPC))

Int Cl' H04N5/92, H04N5/93, G11B20/10, G11B27/10

## B. 調査を行った分野

調査を行った最小限資料(国際特許分類(IPC))

Int Cl' H04N5/92, H04N5/93, G11B20/10, G11B27/10

## 最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922-1996年

日本国公開実用新案公報 1971-2004年

日本国登録実用新案公報 1994-2004年

日本国実用新案登録公報 1996-2004年

国際調査で使用した電子データベース(データベースの名称、調査に使用した用語)

## C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	JP 2003-87744 A (松下電器株式会社) 2003.03.20 全文, 全図 & WO 97/13366 A1 & EP 847200 A1 & US 5884004 A1	1-5, 8, 9-13, 16
A	JP 2002-16880 A (ソニー株式会社) 2002.01.18 全文, 全図 (ファミリーなし)	1, 8, 9, 16
A	JP 2002-199336 A (株式会社東芝) 2002.07.12 段落【0338】 (ファミリーなし)	2-4, 10-12

 C欄の続きにも文献が列挙されている。 パテントファミリーに関する別紙を参照。

## \* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの

「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献(理由を付す)

「O」口頭による開示、使用、展示等に言及する文献

「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&amp;」同一パテントファミリー文献

## 国際調査を完了した日

17.09.2004

## 国際調査報告の発送日

05.10.2004

## 国際調査機関の名称及びあて先

日本国特許庁 (ISA/JP)

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官(権限のある職員)

酒井 朋広

5C 8935

電話番号 03-3581-1101 内線 3541

## 第II欄 請求の範囲の一部の調査ができないときの意見（第1ページの2の続き）

法第8条第3項（PCT17条(2)(a)）の規定により、この国際調査報告は次の理由により請求の範囲の一部について作成しなかった。

1.  請求の範囲 \_\_\_\_\_ は、この国際調査機関が調査をすることを要しない対象に係るものである。つまり、
2.  請求の範囲 6, 7, 14, 15 は、有意義な国際調査をすることができる程度まで所定の要件を満たしていない国際出願の部分に係るものである。つまり、過去に出力された（まだ出力されていない）映像データと関連付けられた識別情報をもつ「出力を停止した音声（データ）」と「出力した音声（データ）」の違いが明確ではない。また、出力を停止した音声と映像とは別々のストリームであるにも係わらず、なぜ関連した識別情報が付与されているのかも不明である。
3.  請求の範囲 \_\_\_\_\_ は、従属請求の範囲であってPCT規則6.4(a)の第2文及び第3文の規定に従って記載されていない。

## 第III欄 発明の単一性が欠如しているときの意見（第1ページの3の続き）

次に述べるようにこの国際出願に二以上の発明があるとこの国際調査機関は認めた。

1.  出願人が必要な追加調査手数料をすべて期間内に納付したので、この国際調査報告は、すべての調査可能な請求の範囲について作成した。
2.  追加調査手数料を要求するまでもなく、すべての調査可能な請求の範囲について調査することができたので、追加調査手数料の納付を求めなかった。
3.  出願人が必要な追加調査手数料を一部のみしか期間内に納付しなかったので、この国際調査報告は、手数料の納付のあった次の請求の範囲のみについて作成した。
4.  出願人が必要な追加調査手数料を期間内に納付しなかったので、この国際調査報告は、請求の範囲の最初に記載されている発明に係る次の請求の範囲について作成した。

## 追加調査手数料の異議の申立てに関する注意

- 追加調査手数料の納付と共に出願人から異議申立てがあった。
- 追加調査手数料の納付と共に出願人から異議申立てがなかった。